Robust VaR and CVaR optimization under joint ambiguity in distributions, means, and covariances

Somayyeh Lotfi · Stavros A. Zenios

To appear: European Journal of Operational Research
Working Paper 16–03
The Wharton Financial Institutions Center
The Wharton School, University of Pennsylvania, PA.
February 2017. Revised August 2017, November 2017, January 2018.

Somayyeh Lotfi
lotfinoghabi–.somayyeh@ucy.ac.cy
Department of Accounting and Finance, Faculty of Economics and Management,
University of Cyprus, P.O. Box 20537, Nicosia, Cyprus.

Stavros A. Zenios (corresponding author)
zenios.stavros@ucy.ac.cy
Department of Accounting and Finance, Faculty of Economics and Management,
University of Cyprus, P.O. Box 20537, Nicosia, Cyprus; Wharton Financial Institutions Center, University of Pennsylvania, Philadelphia, USA; Norwegian School of Economics.

Abstract We develop robust models for optimization of the VaR (value at risk) and CVaR (conditional value at risk) risk measures with a minimum expected return constraint under joint ambiguity in distribution, mean returns, and covariance matrix. We formulate models for ellipsoidal, polytopic, and interval ambiguity sets of the means and covariances. The models unify and/or extend several existing models. We also show how to overcome the well-known conservativeness of robust optimization models by proposing an algorithm and a heuristic for constructing joint ellipsoidal ambiguity sets from point estimates given by multiple securities analysts. Using a controlled experiment we show how the well-known sensitivity of CVaR to mis-specifications of the first four moments of the distribution is alleviated with the robust models. Finally, applying the model to the active management of portfolios of sovereign credit default swaps (CDS) from Eurozone core and periphery, and Central, Eastern and South-Eastern Europe countries, we illustrate that investment strategies using robust optimization models perform well out-of-sample, even during the eurozone crisis. We consider both buy-and-hold and active management strategies.

Keywords Risk management · Data ambiguity · Coherent risk measures · Portfolio optimization · Eurozone crisis.

Contents

Acknowledgements

Stavros Zenios is holder of a Marie Sklodowska-Curie fellowship. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 655092. An early version of the paper benefited from comments by two anonymous referees.

1 Introduction

Mean-variance portfolio optimization from the seminal thesis of Harry Markowitz provided the basis for a descriptive theory of portfolio choice: how investors make decisions. This led to further research in financial economics, with the development of a theory on price formation for financial assets (by William Sharpe) and on corporate finance taxation, bankruptcy and dividend policies (by Merton Miller). These descriptive contributions of the behavior of financial agents were recognized by a joint Nobel Prize in 1990. The prescriptive part of the theory — how investors should make decisions— was also acclaimed by practitioners, and mean-variance models proliferated. Here, however, problems surfaced: mean-variance portfolio optimization is sensitive to perturbations of input data (Best and Grauer, 1991; Chopra and Ziemba, 1993). Since the estimation of market parameters is error prone, the models are severely handicapped. In theory they produce well diversified portfolios but in practice they generate portfolios biased towards estimation errors.

With advances in financial engineering, variance was replaced by more sophisticated risk measures. Value-at-risk (VaR) became an industry standard and written into the Basel II accords to calculate capital adequacy, or calculate insurance premia, or set margin requirements. However, value-at-risk is criticized for being non-convex and it is also computationally intractable to optimize. In a seminal paper Artzner et al. (1999) provided an axiomatic characterization of *coherent* risk measures, and conditional value-at-risk (CVaR) emerged as one such measure. CVaR rose to prominence with the work of Rockafellar and Uryasev (2000) who showed that it can be minimized as a linear program. CVaR optimization emerged as a credible successor to mean-variance models: it is coherent, computationally tractable, and found numerous applications (Zenios, 2007, ch. 5). Basel III shifted from VaR to an expected shortfall measure of risk, to capture tail risk, especially during periods of financial market stress.

With the increased attention placed on CVaR an argument against its use and in favor of VaR also surfaced. VaR estimated from a set of sampled scenarios is a robust statistic, i.e., it is insensitive to small deviations of the underlying distribution from the observed distribution, whereas CVaR is not. Kou et al. (2013) argue that risk measures should be robust but coherent risk measures are not and CVaR lacks a key property.

This paper contributes to an extensive body of literature that seeks to eliminate the sensitivity of CVaR by incorporating data ambiguity in the optimization model. We develop robust models for VaR and CVaR optimization under general ellipsoidal ambiguity sets and for joint ambiguity in means and covariances. These models touch on previous works by others, reviewed below, and extend and/or unify some of the previous contributions. We also develop an algorithm and a heuristic to construct an ellipsoid ambiguity set from point estimates given by multiple securities analysts, and to control the ambiguity set to avoid too conservative solutions. This contribution is, to the best of our knowledge, novel in the literature. We use numerical experiments to highlighting that the models can be robust under first, second, and higher-moment ambiguity, and also that robustness does not necessarily come by sacrificing portfolio performance.

1.1 Review of robust VaR and CVaR models

The 2008 global crisis revived the work of Chicago economist Frank H. Knight (1921) that considers financial and economic data as ambiguous instead of uncertain. Under uncertainty a probability model is known but the random variables are observed with some measurement error, whereas under ambiguity the probability model is unknown. Hence, data mis-specification is not only due to measurement error, that can be reduced with improved estimation techniques, but is an integral part of financial decision making. Data ambiguity deserves attention as an issue to be modeled, not as a problem to be eliminated. It is from this perspective that we develop this study.

Early suggestions in dealing with the sensitivity of portfolio optimization models to data estimation errors use Bayesian or James-Stein estimators, resampling, or restricting portfolio choices with ad hoc constraints. We do not review this literature as it is outside the scope of our work. We build, instead, on recent research that brings developments in *robust optimization* to bear on portfolio selection under data ambiguity.

Robust optimization models require constraints to be satisfied even with ambiguous data, and the objective value to be insensitive to the ambiguity. Concepts of robustness in optimization have been developed independently in the fields of operations research and engineering design. Mulvey et al. (1995) proposed the robust optimization of large scale systems when data take values from a discrete scenario set, using a regularization of the objective function to control its sensitivity, and penalty functions to control constraint violation. This approach spurred numerous applications in facility location, power capacity planning, disaster response, agribusiness, supply chain management, production and process planning, network design, and so on. Robust convex optimization was developed by Ben-Tal and Nemirovski (1998) for optimization problems with data ambiguity described by an ellipsoid. They showed that important convex optimization problems admit a tractable robust counterpart. The foundational paper spurred extensive theoretical and applied research (Ben-Tal et al., 2009; Bertsimas et al., 2011). In a way, robust portfolio optimization brings ideas from Taguchi robust engineering design to the design of portfolios. Authors usually adopt the robust convex optimization framework over an appropriate ambiguity set, and it is in this domain that our paper makes a contribution. Fabozzi et al. (2010) review robust portfolio optimization using mean, VaR, and CVaR risk measures.

The robust counterpart to mean-variance optimization was developed by Goldfarb and Iyengar (2003). Using a linear factor model for asset returns they introduce "uncertainty structures" —the confidence regions associated with parameter estimation— and formulate robust portfolio selection models corresponding to these uncertainty structures as second order cone programs (SOCP). They also develop robust counterparts for VaR and CVaR optimization under the normality assumption of mean-variance models. Schottle and Werner (2009); Tütüncü and Koenig (2004); Ye et al. (2012) develop further robust mean-standard deviation and mean-variance models, removing some of the assumptions of the Goldfarb–Iyengar paper.

Our paper develops a robust counterpart of CVaR optimization (RCVaR) and finds it identical to robust VaR optimization (RVaR). Hence, we give a detailed review of previous works on RVaR and RCVaR optimization, so we can place the innovations of our own contribution.

Current literature addresses the following sources of ambiguity of model parameters: (i) ambiguity in mean return estimates, (ii) ambiguity in covariance matrix estimates, and (iii) ambiguity in the distribution of the data. Ambiguity can be independent for each parameter or joint for multiple parameters. If ambiguity is independent for each parameter we have simple sets of parameter values, e.g., a (sub)vector of parameters lies in some interval. For joint ambiguity the parameters belong to sets such as ellipsoids or convex polytopes. Models based on discrete scenarios may have ambiguity in the scenario values or the scenario probabilities or both. For models with continuous distributions, ambiguity is in the moments. Another important distinction in understanding the various contributions is captured in the terminology used. VaR and CVaR *minimization* models lack a minimum return constraint, and ambiguity is restricted to the objective function. VaR and CVaR *optimization* trade off the risk measure against an expected return target. The difference between risk-minimization and risk-optimization is not innocuous: risk-optimization models with a target expected return are difficult to analyze if the means are ambiguous, as ambiguity appears in the constraints, and not only in the obective.

El Ghaoui et al. (2003) address RVaR minimization with partially known distributions of returns, whereby means and covariance lie within a known uncertainty set, such as an interval, a polytope (polytopic uncertainty), or a convex subset (convex moment uncertainty). Given this information on return distributions they cast RVaR minimization for interval uncertainty as semidefinite program (SDP), and for polytopic uncertainty as SOCP. They also give a general, but potentially intractable, model for convex moment uncertainty. Their model lacks the target expected return constraint.

The first RCVaR optimization model is by Quaranta and Zaffaroni (2008) for interval uncertainty of the means. Zhu and Fukushima (2009) consider RCVaR optimization for box and ellipsoidal uncertainty in distribution, as well as distribution mixtures of convex combination of predetermined distributions and unknown mixture weights. By "distribution" the authors mean the probabilities of the discretized data. Their model can potentially be used to approximate joint ambiguity of means, covariances, and higher moments through the choice of weights of the distributions, using results of Marron and Wand (1992)[1]. Distinctly from this work, our models are exact. Furthermore, if there is no information on the distribution of the discretized random variable but there is information on the first two moments of the distribution, then we can use our models but not Zhu-Fukushima.

Delage and Ye (2010) show (as a special case of their work) that RCVaR minimization for ambiguity in the probabilities, mean, and second moment, can be solved in polynomial time. The authors provide bounds and generate confidence regions on the mean and covariance matrix in case of moment uncertainty but stop short from developing RCVaR models and develop, instead, a robust model for expected utility maximization under moment uncertainty.

Chen et al. (2011) point out that robust solutions come with a computational price: robust optimization models can be infinite dimensional and, without proper choice of uncertainty sets, the model may be intractable. They obtain bounds on worst case value of lower partial moments and use them to develop RVaR and RCVaR minimization for distribution ambiguity with closed form solution under a normalization constraint. Paç and Pınar (2014) extend further RVaR and RCVaR optimization for distribution and mean returns ambiguity, but fixed covariance matrix. This is one of the papers extended in our work, by allowing covariance matrix ambiguity.

A contribution that filled several gaps is Gotoh et al. (2013). Scenario based VaR and CVaR minimization models use discrete data observations (i.e., scenarios) and their probabilities to determine the empirical distribution. There are three possible ways to introduce ambiguity and formulate RVaR and RCVaR counterparts using scenarios. The first approach (Zhu and Fukushima, 2009), keeps the scenarios fixed and considers ambiguous probabilities from a box or an ellipsoid. Gotoh et al. (2013) consider a second approach with uncertainty in scenarios but fixed probabilities, and a third approach, where both scenarios and probabilities are ambiguous.

Our work considers ambiguity in the distribution as well as mean returns and the covariance matrix, and joint ambiguity in combinations of the above. These are, to the best of our knowledge, the most general ambiguity sets considered in the literature for RVaR and RCVaR models. Joint uncertainty in means and covariance matrix was also considered by Schottle and Werner (2009) and Ye et al. (2012) but for mean-standard deviation and mean-variance optimization, respectively. We use an ellipsoidal ambiguity set which is general and obtain tractable optimization models as SOCP. We use the term ambiguity sets in the Knightean sense, instead of uncertainty sets, in discussing robust models. Robust optimization literature typically refers to uncertainty sets although usually ambiguity is meant.

The paper is organized as follows. Section 2 defines VaR, CVaR, RVaR, and RCVaR models and discusses the instability of VaR and CVaR optimal portfolios. Section 3 is the main one. It formulates RVaR and RCVaR for ambiguous distributions, and ellipsoidal ambiguity in means and covariance, discusses the construction of ambiguity sets (sec. 3.2), extends or unifies existing results (sec. 3.3). It also develops models for polytopic and interval ambiguity sets (sec. 3.4). In developing our models we also identify some implicit limiting assumptions made in previous works and explain how we overcome the limitations. Section 4 reports on two distinct numerical tests. First, using simulations we test the robustness of optimal portfolios

[1] We thank a referee for pointing out this potential generalization.

under mis-specification of mean, variance, skewness, and kurtosis. Second, using sovereign CDS spread returns during the eurozone crisis we investigate the robustness of alternative investment strategies. Conclusions are in section 5. Proofs are gathered in the appendix.

2 Optimization of VaR and CVaR risk measures and their stability

The mean of α-tail[2] distribution of the loss random variable Ξ, $\mathrm{CVaR}_\alpha(\Xi)$, and its minimization formula are given by the following theorem of Rockafellar and Uryasev (2000).

Theorem 1 ***Fundamental minimization formula.***
As a function of $\gamma \in \mathbb{R}$, the auxiliary function

$$F_\alpha(\Xi,\gamma) = \gamma + \frac{1}{1-\alpha}\mathbb{E}\{[\Xi-\gamma]^+\},$$

where $\alpha \in (0,1]$ is the confidence level and $[t]^+ = max\{0,t\}$, is finite and convex, with

$$CVaR_\alpha(\Xi) = \min_{\gamma\in\mathbb{R}} F_\alpha(\Xi,\gamma).$$

Moreover, the set M_α of minimizers to $F_\alpha(\Xi,\gamma)$ is a compact interval, $M_\alpha = [x_\alpha, x^\alpha]$, where $x_\alpha = inf\{x \in \mathbb{R} : P[\Xi \le x] \ge \alpha\}$ and $x^\alpha = inf\{x \in \mathbb{R} : P[\Xi \le x] > \alpha\}$.

Remark 1 Note that x_α, the left end-point of the set M_α, and not every minimizer of $F_\alpha(\Xi,\gamma)$, is equal to $\mathrm{VaR}_\alpha(\Xi)$. Hence, the statement $\mathrm{VaR}_\alpha(\Xi) = \operatorname*{argmin}_{\gamma\in\mathbb{R}} F_\alpha(\Xi,\gamma)$ is true only when the minimum is unique and the interval reduces to a point.

Consider an investor operating in a market with n risky assets, a riskless asset, and no short-selling. The riskless asset has rate of return r_f and the n risky assets have rates of return denoted by random vector ξ. The loss function associated with decision variable $x \in \mathbb{R}^n$ of proportionate allocations to the risky assets is given by

$$f(x,\xi) = -(x^\top\xi + r_f(1 - x^\top e)),$$

where e is an n-vector of ones. (When dealing with portfolio optimization models, loss is a function of the portfolio x, i.e., the loss random variable is $\Xi \doteq \Xi(x)$, and we write the auxiliary function and CVaR as functions of x.) According to Theorem 1 the conditional value-at-risk of the loss function is the solution of

$$\mathrm{CVaR}_\alpha(x) = \min_{\gamma\in\mathbb{R}} F_\alpha(x,\gamma), \tag{1}$$

where

$$F_\alpha(x,\gamma) = \gamma + \frac{1}{1-\alpha}\mathbb{E}\{[f(x,\xi)-\gamma]^+\}.$$

If $\bar{\gamma}$ denotes $\operatorname*{argmin}_{\gamma} F_\alpha(x,\gamma)$, then, by Theorem 1, $\mathrm{VaR}_\alpha(x)$ is obtained from

$$\begin{aligned} \mathrm{VaR}_\alpha(x) \;=\; & \min_{\gamma\in\mathbb{R}} \quad \gamma \\ & s.t.\ F_\alpha(x,\gamma) \le F_\alpha(x,\bar{\gamma}). \end{aligned} \tag{2}$$

This problem is non-convex, unless $F_\alpha(x,\gamma)$ has a unique minimum, e.g., for normally distributed returns with strictly increasing cumulative distribution function.

[2] We use $\alpha = 0.95$ in all numerical experiments throughout the paper.

The definition of VaR in (2) uses the auxiliary function $F_\alpha(x,\gamma)$, whereas the original formulation of VaR is

$$\begin{aligned} \text{VaR}_\alpha(x) \;=\; & \min_{\gamma\in\mathbb{R}} \;\; \gamma \\ & s.t.\ \text{Prob}\,\{\gamma \le f(x,\xi)\} \le 1-\alpha. \end{aligned} \tag{3}$$

Hence, models for selecting a portfolio with minimal VaR or CVaR, and a minimum return constraint, can be posed as follows:

I. VaR optimization

$$\begin{aligned} \min_{\gamma\in\mathbb{R},\, x\in\mathbb{R}^n} \quad & \gamma \\ s.t. \quad & F_\alpha(x,\gamma) \le F_\alpha(x,\bar{\gamma}), \\ & (\bar{\mu} - r_f e)^\top x \ge d - r_f, \end{aligned} \tag{4}$$

or

$$\begin{aligned} \min_{\gamma\in\mathbb{R},\, x\in\mathbb{R}^n} \quad & \gamma \\ s.t. \quad & \text{Prob}\,\{\gamma \le f(x,\xi)\} \le 1-\alpha, \\ & (\bar{\mu} - r_f e)^\top x \ge d - r_f. \end{aligned} \tag{5}$$

II. CVaR optimization

$$\begin{aligned} \min_{\gamma\in\mathbb{R}, x\in\mathbb{R}^n} \quad & F_\alpha(x,\gamma) \\ s.t. \quad & (\bar{\mu} - r_f e)^\top x \ge d - r_f. \end{aligned} \tag{6}$$

$\bar{\mu}$ is the risky assets mean vector and $d \in \mathbb{R}^+$ is the minimum return satisfying $d \ge r_f$.

Without loss of generality, in our numerical work we will consider VaR and CVaR minimization with a budget constraint, no risk-free asset, and no short-selling. Denote by $\mathbb{X}$ the set $\{x \in \mathbb{R}^n \mid x \ge 0,\, \sum_{i=1}^n x_i = 1\}$, and use Monte Carlo simulation to generate an $S \times n$ matrix R of return scenarios for n risky assets, where S is the number of scenarios generated by the simulations. Using M to denote a large positive number, VaR model (5) is solved as the mixed integer linear program, see, e.g., Pflug (2000),

$$\begin{aligned} \min_{x\in\mathbb{X},\, \gamma\in\mathbb{R},\, y\in\{0,1\}^S} \quad & \gamma \\ s.t. \quad & \\ & -Rx - My - e\gamma \le 0, \\ & e^\top y \le S(1-\alpha). \end{aligned} \tag{7}$$

CVaR model (6) is formulated as the linear program (Rockafellar and Uryasev, 2000)

$$\begin{aligned} \min_{x\in\mathbb{X},\, u\in\mathbb{R}^S,\, \gamma\in\mathbb{R}} \quad & \gamma + \frac{1}{S(1-\alpha)} e^\top u \\ s.t. \quad & \\ & -Rx - e\gamma \le u, \\ & u \ge 0. \end{aligned} \tag{8}$$

It is well-known that scenario based CVaR is not a robust estimator whereas VaR is, and Gotoh et al. (2013); Lim et al. (2011) also noted the instability of the portfolio weights obtained when minimizing these measures. This instability is exemplified in Figure 11, where it is also shown that the robust models generate more stable portfolios.

3 Robust VaR and CVaR for distribution and moment ambiguity

We introduce now ambiguity in the models. Robust counterparts for VaR and CVaR are formulated as SOCPs and we will observe that they are identical. We consider a joint ellipsoidal structure for the ambiguity set of mean returns and covariance matrix. Ellipsoidal sets can be viewed as generalizations of polytopic sets (Ben-Tal et al., 2009), and therefore our model generalizes El Ghaoui et al. (2003) who modeled RVaR minimization for polytopic uncertainty. Extending the models from sets with independence between means and covariance (El Ghaoui et al., 2003; Goldfarb and Iyengar, 2003; Tütüncü and Koenig, 2004) to sets that capture dependencies, we generate better diversified and less conservative portfolios as argued by Lu (2011) for mean-variance models.

Definition 1 (Ambiguity in distribution) The random variable ξ assumes a distribution from

$$\mathbb{D} = \{\pi \mid \mathbb{E}_\pi[\xi] = \bar{\mu},\, \mathrm{Cov}_\pi[\xi] = \bar{\Gamma} \succ 0\},$$

where $\bar{\mu}$ and $\bar{\Gamma}$ are given.

Definition 2 (Ellipsoidal ambiguity for mean returns and covariance matrix) Mean returns and covariance matrix belong to the joint ellipsoidal set:

$$U_\delta(\hat{\mu},\hat{\Gamma}) = \{(\bar{\mu},\bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid S(\bar{\mu}-\hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}-\hat{\mu}) + \frac{S-1}{2}\|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}-\hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2 \le \delta^2\},$$

where $\|A\|_{tr}^2 = tr(AA^\top)$. (For the definition see (Schottle and Werner, 2009, Proposition 3.3).)

Remark 2 The set $U_\delta(\hat{\mu},\hat{\Gamma})$ is a generalization of ambiguity sets that have been used in the literature. Setting $\bar{\Gamma} = \hat{\Gamma}$, i.e., certainty about the estimate of the covariance matrix, we obtain the ellipsoidal set for mean returns used in Ceria and Stubbs (2006); Chen et al. (2011); Schottle and Werner (2009); Zhu et al. (2008). Similarly, if we fix $\bar{\mu} = \hat{\mu}$, i.e., certainty about the mean estimates, we obtain the ellipsoidal set for covariance matrix. Goldfarb and Iyengar (2003) use this structure of uncertainty set for the factor loading matrix of a factor model of returns.

Remark 3 $U_\delta(\hat{\mu},\hat{\Gamma})$ can be decomposed to $U_{\sqrt{\kappa}\delta}(\hat{\mu})$ and $U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma})$ using a parameter $\kappa \in [0,1]$, with $U_{\sqrt{\kappa}\delta}(\hat{\mu}) = \{\bar{\mu} \in \mathbb{R}^n \mid S(\bar{\mu}-\hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}-\hat{\mu}) \le \kappa\delta^2\}$ and $U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma}) = \{\bar{\Gamma} \in \mathbb{S}^n \mid \frac{S-1}{2}\|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}-\hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2 \le (1-\kappa)\delta^2\}$. This representation is used later to obtain, as special cases, models with ambiguity in means ($\kappa = 1$) or covariances ($\kappa = 0$) only.

To develop RVaR and RCVaR models we start from one of the following:

1. $\mathrm{RVaR_I}$ (robust counterpart of model (4))

$$\begin{aligned} \min_{\gamma\in\mathbb{R},\, x\in\mathbb{R}^n} \quad & \gamma \\ s.t. \quad & \max_{(\bar{\mu},\bar{\Gamma})\in U_\delta(\hat{\mu},\hat{\Gamma}),\, \pi\in\mathbb{D}} [F_\alpha(x,\gamma) - F_\alpha(x,\bar{\gamma})] \le 0, \\ & \min_{(\bar{\mu},\bar{\Gamma})\in U_\delta(\hat{\mu},\hat{\Gamma}),\, \pi\in\mathbb{D}} (\bar{\mu} - r_f e)^\top x \ge d - r_f. \end{aligned} \tag{9}$$

2. $\mathrm{RVaR_{II}}$ (robust counterpart of model (5))

$$\begin{aligned} \min_{\gamma\in\mathbb{R},\, x\in\mathbb{R}^n} \quad & \gamma \\ s.t. \quad & \max_{(\bar{\mu},\bar{\Gamma})\in U_\delta(\hat{\mu},\hat{\Gamma}),\, \pi\in\mathbb{D}} \mathrm{Prob}\{\gamma \le f(x,\xi)\} \le 1-\alpha, \\ & \min_{(\bar{\mu},\bar{\Gamma})\in U_\delta(\hat{\mu},\hat{\Gamma}),\, \pi\in\mathbb{D}} (\bar{\mu} - r_f e)^\top x \ge d - r_f, \end{aligned} \tag{10}$$

where $\bar{\gamma} = \mathrm{argmin}_\gamma F_\alpha(x,\gamma)$.

3. RCVaR (robust counterpart of model (6))

$$\begin{aligned} &\min_{x\in\mathbb{R}^n,\,\gamma\in\mathbb{R}} \max_{(\bar\mu,\bar\Gamma)\in U_\delta(\hat\mu,\hat\Gamma),\,\pi\in\mathbb{D}} F_\alpha(x,\gamma) \\ &s.t. \min_{(\bar\mu,\bar\Gamma)\in U_\delta(\hat\mu,\hat\Gamma),\,\pi\in\mathbb{D}} (\bar\mu - r_f e)^\top x \geq d - r_f. \end{aligned} \tag{11}$$

Remark 4 The maximization problem in the first constraint of (9) can not, in general, be solved explicitly. Existing papers for RVaR minimization (Chen et al., 2011) and RVaR optimization (Paç and Pınar, 2014) solve the special case when $F_\alpha(x,\gamma)$ has a unique minimum[3], thereby obtaining an explicit solution to the maximization problem in the constraint and simplifying the RVaR formulation.

Remark 5 Unique minimum of $F_\alpha(x,\gamma)$ implies unique solution of $\text{Prob}\{f(x,\xi)\leq\gamma\}=\alpha$, which is then $\text{VaR}_\alpha(x)$. This occurs when the distribution function of portfolio loss is strictly increasing. However, the loss distribution function is often a (non-decreasing) continuous step function, and Rockafellar and Uryasev (2002) extended their original contribution to derive the fundamental properties of CVaR for general loss distributions. We work with (10) to deal with the inner maximization for general loss distributions.

3.1 Explicit formulation of RVaR and RCVaR optimization models

We obtain now explicit formulations for models (10) and (11). First we prove an essential proposition and then the main theorem.

Proposition 1 *If random variable ξ has a distribution from the set $\mathbb{D}$ with fixed $\bar\mu$ and $\bar\Gamma$, then*

$$\begin{aligned} \min_{\gamma\in\mathbb{R}} \max_{\pi\in\mathbb{D}} F_\alpha(x,\gamma) = &\min_{\gamma\in\mathbb{R},\,x\in\mathbb{R}^n} \gamma \\ &s.t. \max_{\pi\in\mathbb{D}} Prob\{\gamma\leq f(x,\xi)\}\leq 1-\alpha, \end{aligned} \tag{12}$$

and the solution is $-r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}$.

Proof. *From equation (10) in Paç and Pınar (2014) we know that*

$$\min_{\gamma\in\mathbb{R}} \max_{\pi\in\mathbb{D}} F_\alpha(x,\gamma) = -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}. \tag{13}$$

For the constraint of the optimization problem in the right-hand side of (12), we use Theorem 1 of El Ghaoui et al. (2003) (set $1-\alpha$ and $f(x,\xi)$ instead of ϵ and $-r(w,x)$, respectively), which means

$$\max_{\pi\in\mathbb{D}} Prob\{\gamma\leq f(x,\xi)\}\leq 1-\alpha$$

is equivalent to

$$-r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}\leq\gamma.$$

Hence, the optimization problem in the right-hand side is equivalent to

$$\begin{aligned} &\min_{\gamma\in\mathbb{R}} \gamma \\ &s.t. -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}\leq\gamma, \end{aligned} \tag{14}$$

[3] This implicit assumption is made in the proofs of Theorems 2.9 and 1, respectively, when the authors invoke the equality $\text{VaR}_\alpha(\Xi)=\operatorname{argmin}_{\gamma\in\mathbb{R}} F_\alpha(\Xi,\gamma)$ which holds true only when $F_\alpha(x,\gamma)$ has a unique minimum, see Remark 1.

which has the minimum value $-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top \bar{\Gamma} x}$. *This completes the proof.*

Remark 6 The left and right-hand sides of (12) are RCVaR and RVaR, respectively, associated with the ambiguity set of Definition 1. Hence, the robust counterpart to VaR and CVaR optimization for distribution ambiguity is the same optimization model.

Zymler et al. (2013a,b) also derived worst-case VaR and CVaR for piece-wise and quadratic loss functions when there is ambiguity in distribution of asset returns but known mean and covariance information. Both of optimization problems are obtained as SDPs and they turned out to be identical for both loss functions. When the loss function is linear (no derivative securities), the associated problems reduce to the same SOCPs as in our Proposition 1. Zymler et al. (2013a) extend further these identical models to ambiguity in means and covariance matrix. They take into account box-type ambiguity in the moment matrix but we consider joint ellipsoidal ambiguity set for mean and covariance matrix and obtain the robust models for a linear loss function next. So, compared to Zymler et al. we establish the equivalence with more general ambiguity sets when there is ambiguity in means and covariance matrix, whereas they use more general loss functions. Our result also generalizes the result of Čerbáková (2006) for the special case of symmetric distributions identified only by the first two moments. It also establishes that the bounds obtained by Bertsimas et al. (2004) on VaR and CVaR for distribution ambiguity are tight.

We obtain now RVaR and RCVaR optimization models for ambiguity in distributions, mean returns and covariance matrix.

Theorem 2 *If random variable* ξ *has a distribution from the set* $\mathbb{D}$ *and* $(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma})$. *Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) are both represented by the following SOCP:*

$$\begin{aligned} \min_{x\in\mathbb{R}^n} & - r_f - (\hat{\mu} - r_f e)^\top x + \left(\max_{\kappa\in[0,1]} f(\kappa)\right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \\ s.t. & - \frac{\delta}{\sqrt{S}}\|\hat{\Gamma}^{\frac{1}{2}} x\| + (\hat{\mu} - r_f e)^\top x \geq d - r_f, \end{aligned} \tag{15}$$

where $f(\kappa) = \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{(1+\delta\sqrt{\frac{2(1-\kappa)}{S-1}}} + \delta\sqrt{\frac{\kappa}{S}}$.
Proof. *See Appendix A.1.*

Remark 7 $f(\kappa)$ is a strictly concave function with $lim_{\kappa\to 0} f^{'}(\kappa) = \infty$ and $lim_{\kappa\to 1} f^{'}(\kappa) = -\infty$. Hence, $f(\kappa)$ has a unique maximum in the interval $(0, 1)$.

3.2 Constructing the ambiguity set

The ambiguity set is typically taken as input in the robust optimization literature. In some cases the ambiguity set can be defined as the confidence regions of the statistical estimators of the model parameters (Goldfarb and Iyengar, 2003; Schottle and Werner, 2009). For other cases —such as in the applications we solve later— we may be given multiple estimates of model parameters which raises the question as to what is the appropriate ambiguity set. In finance, for instance, it is not uncommon to be given estimates by multiple securities analysts and we need a method to construct an ambiguity set, including its center. This issue has not been addressed in existing literature and we now propose and solve analytically a nonlinear SDP for finding the center of a joint ellipsoidal set.

Assume K experts provide estimates for mean returns and covariance matrices $(\bar{\mu}_k, \bar{\Gamma}_k)$, $k = 1, 2, \ldots, K$. (For convenience we assume they were all estimated using the same number of scenarios S.) To construct their joint ellipsoidal ambiguity set we need to fix the center ($\hat{\mu}$,

$\hat{\Gamma}$). This is obtained as the solution of a nonlinear convex program for minimizing the l_2-norm of the parameters δ_k, for $k = 1, 2, \ldots, K$, where each parameter corresponds to the ellipsoid with center $(\hat{\mu}, \hat{\Gamma})$ containing observation $(\bar{\mu}_k, \bar{\Gamma}_k)$. Referring to Definition 2 the optimization problem is given by

$$\min_{\hat{\mu}\in\mathbb{R}^n,\, \hat{\Gamma}\in\mathbb{S}^n_{++}} \sqrt{\sum_{k=1}^{K} S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2}\|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|^2_{tr}}, \tag{16}$$

where $\mathbb{S}^n_{++}$ is the set of all n-dimensional, symmetric, positive definite matrices. This problem is equivalent to

$$\min_{\hat{\mu}\in\mathbb{R}^n,\, \hat{\Gamma}\in\mathbb{S}^n_{++}} \sum_{k=1}^{K} S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2}\|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|^2_{tr}. \tag{17}$$

The next theorem gives the solution of this problem, if a solution exists.

Theorem 3 *If (17) is solvable, then it admits the following solution:*

1. $\hat{\mu} = \frac{1}{K}\sum_{k=1}^{K} \bar{\mu}$.
2. Γ^-, *the inverse of optimal* $\hat{\Gamma}$, *is obtained from the linear system of equations*

$$\left[\sum_{k=1}^{K} \bar{\Gamma}_k \otimes \bar{\Gamma}_k\right] vec(\Gamma^-) = \sum_{k=1}^{K} vec(\bar{\Gamma}_k) - \frac{S}{(S-1)}\sum_{k=1}^{K} vec\left((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top\right), \tag{18}$$

where $\otimes$ *is the Kronecker product and* $\sum_{k=1}^{K} \bar{\Gamma}_k \otimes \bar{\Gamma}_k$ *is positive semidefinite.*

If at least one of $\bar{\Gamma}_k$, $k = 1, \ldots, K$, *is positive definite, then (18) has a unique solution.*
Proof. *See Appendix A.2.*

We now state a simple algorithm for constructing ellipsoidal ambiguity sets.

3.2.1 Algorithm for constructing a joint ellipsoidal ambiguity set

1. Compute $\hat{\mu} = \frac{1}{K}\sum_{k=1}^{K} \bar{\mu}$, solve the system of linear equations (18) for $vec(\Gamma^-)$, form matrix Γ^-, and calculate its inverse $\hat{\Gamma}$ to obtain the center $(\hat{\mu}, \hat{\Gamma})$.
2. Choose δ such that the resulting ellipsoidal set inscribes $(\bar{\mu}_k, \bar{\Gamma}_k)$, $k = 1, \ldots, K$, i.e., compute the distance of each estimate from the center, $\delta_1, \ldots, \delta_K$, and let δ be the maximum value.

System (18) is of dimension $n^2 \times n^2$ which depends only on the number of assets n. Methods for solving systems of equations based on LU or Cholesky (when applicable) factorizations are polynomial of cubic order. Hence, the computational complexity of the algorithm is $O(n^6)$, and for medium portfolio sizes this is tractable.

3.2.2 Heuristic for constructing a joint ellipsoidal ambiguity set

It is also possible to construct the ambiguity set with a simple heuristic. In some cases the heuristic gives tighter ellipsoids than the algorithm and this results to less conservative robust solutions. The heuristic needs K inversions of a matrix of dimension $n \times n$ and elementary matrix operations, and its computational complexity is $O(Kn^3)$.

1. For each estimate $(\bar{\mu}_k,\, \bar{\Gamma}_k)$ we compute the sum of its distances from all others

$$\text{dist}_k = \sqrt{\sum_{k'=1}^{K} S(\bar{\mu}_{k'} - \bar{\mu}_k)^\top \bar{\Gamma}_k^{-1}(\bar{\mu}_{k'} - \bar{\mu}_k) + \frac{S-1}{2}\|\bar{\Gamma}_k^{-\frac{1}{2}}(\bar{\Gamma}_{k'} - \bar{\Gamma}_k)\bar{\Gamma}_k^{-\frac{1}{2}}\|^2_{tr}}.$$

2. The estimate with the minimum value of dist_k is the center, and we choose δ such that the constructed ellipsoidal set inscribes all $(\bar{\mu}_k, \bar{\Gamma}_k)$, $k = 1, \ldots, K$. Specifically, we compute $\delta_1, \ldots, \delta_K$ as the distance of each point in the ellipsoid from the center and let δ be the maximum value.

3.2.3 Comments on the choice of method

It remains an open question how to construct an ellipsoidal ambiguity set that is big enough to guarantee robustness but tight enough to avoid conservative solutions. One may wish to try both the algorithm and the heuristic and pick the ellipsoid with the smaller δ, knowing that both ellipsoidal sets ensure robust solutions and the one with the smallest δ is less conservative. Furthermore, both methods provide an intuitive way to choose smaller values of δ by choosing a suitable quantile of δ_k, $k = 1, \ldots, K$. It is not clear on the outset which method generates the tighter ellipsoid. Figure 1 illustrates in two-dimensions situations when one method dominates the other. When observations are evolving slowly, or differ slightly from each other, the heuristic performs better since one of the observations provides a good approximation to the ellipsoid's center. When observations change significantly then the algorithm is better in finding a center of the diverse observations.

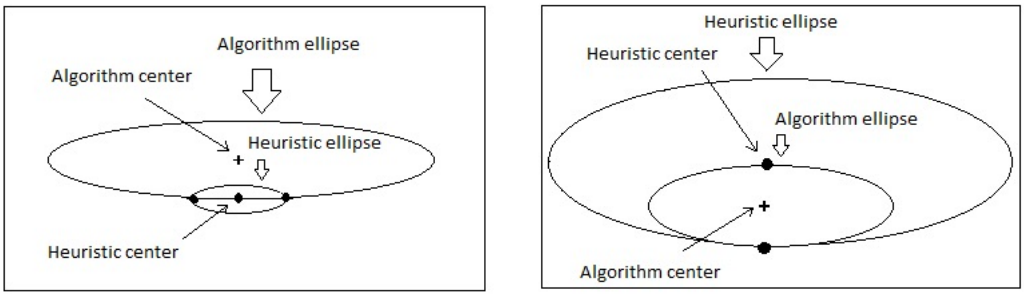


Fig. 1: Illustrating algorithm- and heuristic-constructed ellipsoids in two dimensions. For the observations (solid bullets) on the left, with small changes from each other, the heuristic estimates the tighter ellipsoid. For the observations on the right, with bigger changes, the algorithm ellipsoid is tighter. When there is no ambiguity in (co)variance the centers coincide. The figure illustrates the general case in two dimensions. The x-axis represents mean values and the x-coordinate of the centers coincide. The y-axis representing variances and the y-coordinates are obtained from eqn (18), which is not a simple average, and hence the y-coordinates of the centers do not coincide.

3.3 Unifying and extending some results on RVaR and RCVaR optimization

From our model we obtain, as special cases, known results from the literature.

1. For distribution ambiguity, but known mean returns and covariance matrix, set $\bar{\mu} = \hat{\mu}$, $\bar{\Gamma} = \hat{\Gamma}$ in Proposition 1, to get the results of Chen et al. (2011); Paç and Pınar (2014) for RCVaR, and of El Ghaoui et al. (2003) for RVaR.

2. For ambiguity in distribution and means, but known covariance, set $\kappa = 1$ in Remark 3 to obtain the problem studied by Paç and Pınar (2014). Following the proof of Theorem 2 we

now get

$$\min_{x\in\mathbb{R}^n} - r_f - (\hat{\mu} - r_f e)^\top x + \left(\frac{\delta}{\sqrt{S}} + \sqrt{\frac{\alpha}{1-\alpha}}\right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \tag{19}$$
$$s.t. - \frac{\delta}{\sqrt{S}}\|\hat{\Gamma}^{\frac{1}{2}} x\| + (\hat{\mu} - r_f e)^\top x \geq d - r_f,$$

which is the Paç–Pınar RCVaR model, with their constant $\epsilon = \frac{\delta}{\sqrt{S}}$.

3. For ambiguity in distribution and covariance, but known means, set $\kappa = 1$ in Remark 3 and follow the proof of Theorem 2 to get

$$\min_{x\in\mathbb{R}^n} -r_f - (\hat{\mu} - r_f e)^\top x + \left(\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{(1+\delta\sqrt{\frac{2}{S-1}}}\right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \tag{20}$$
$$s.t. \;\; (\hat{\mu} - r_f e)^\top x \geq d - r_f.$$

We are not aware of any studies of this case, which is of interest for risk minimization or when there is special knowledge on the mean return. For instance, in index tracking (Zenios, 2007, ch. 7) the mean excess return of a portfolio over the index is zero in efficient markets.

Using Theorem 2 we can relax the assumption of Chen et al. (2011); Paç and Pınar (2014), see Remark 4. Their RVaR model is

$$-r_f - (\bar{\mu} - r_f e)^\top x + \frac{2\alpha - 1}{2\sqrt{\alpha}\sqrt{1-\alpha}}\sqrt{x^\top \bar{\Gamma} x},$$

while we (and El Ghaoui et al. (2003)) have

$$-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top \bar{\Gamma} x}.$$

From Remark 5 we know that their RVaR model is computed over a subset of the original ambiguity set, i.e., the set of all *strictly increasing* distribution functions with fixed means and covariance, and their parameter differs from ours. Obviously $\frac{2\alpha-1}{2\sqrt{\alpha}\sqrt{1-\alpha}} < \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}$ and their robust counterpart is less conservative but is valid only under the assumption.

3.4 Extensions to polytopic and interval ambiguity sets

We extend now to models for distribution ambiguity and polytopic and interval ambiguity sets in the mean returns and covariance matrix. Polytopic and interval uncertainty were studied by El Ghaoui et al. (2003) for minimizing RVaR, without a target expected return constrain. Thus, our models extend previous work to include target return constraints for RVaR optimization. By Proposition 1, we also have a RCVaR optimization model for polytopic ambiguity, that, to the best of our knowledge is novel. We also obtain a model for RCVaR optimization for interval ambiguity which is different from that of Zhu and Fukushima (2009), as discussed above.

3.4.1 Extension to polytopic ambiguity sets

We give the formal definition and the relevant theorem to specify RVaR and RCVaR optimization models for polytopic ambiguity sets on the mean returns and covariance matrix.

Definition 3 (Polytopic ambiguity for mean returns and covariance matrix) Mean returns and covariance matrix belong to the following polytopic set:

$$U_P = \{(\bar{\mu}, \bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid \bar{\mu} = \sum_{j=1}^{J} \rho_j \bar{\mu}_j, \bar{\Gamma} = \sum_{j=1}^{J} \rho_j \bar{\Gamma}_j, \sum_{j=1}^{J} \rho_j = 1,\ \rho_j \geq 0,\ j = 1, \ldots, J\},$$

where $(\bar{\mu}_j, \bar{\Gamma}_j) \in \mathbb{R}^n \times \mathbb{S}^n_+$, $j = 1, \ldots, J$, are the polytope vertices.

Remark 8 Similar to El Ghaoui et al. (2003) we consider the case when mean and covariance matrix are subject to independent polytopic ambiguity sets. That is we let U_p be the direct product of two polytopes U_{P_1} and U_{P_2}, where

$$U_{P_1} = \{\bar{\mu} \in \mathbb{R}^n \mid \bar{\mu} = \sum_{j=1}^{J} \rho_j \bar{\mu}_j, \sum_{j=1}^{J} \rho_j = 1,\ \rho_j \geq 0,\ j = 1, \ldots, J\}$$

$$U_{P_2} = \{\bar{\Gamma} \in \mathbb{S}^n \mid \bar{\Gamma} = \sum_{j=1}^{J} \rho_j \bar{\Gamma}_j, \sum_{j=1}^{J} \rho_j = 1,\ \rho_j \geq 0,\ j = 1, \ldots, J\}$$

are polytopic ambiguity sets for means and covariance matrix, respectively.

Theorem 4 *If random variable ξ has a distribution from the set $\mathbb{D}$ and $(\bar{\mu}, \bar{\Gamma}) \in U_P$. Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) for polytopic ambiguity are represented by the following SOCP:*

$$\begin{aligned} \min_{x \in \mathbb{R}^n,\, \omega \in \mathbb{R},\, \beta \in \mathbb{R}} \quad & \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \omega - \beta \\ s.t. \quad & \\ & r_f + (\bar{\mu}_j - r_f e)^\top x \geq d, \quad j = 1, \ldots, J, \\ & r_f + (\bar{\mu}_j - r_f e)^\top x \geq \beta, \quad j = 1, \ldots, J, \\ & \sqrt{x^\top \bar{\Gamma}_j x} \leq \omega, \quad j = 1, \ldots, J. \end{aligned} \tag{21}$$

Proof. *See Appendix A.3*

3.4.2 Extension to interval ambiguity sets

We give now the formal definition and the relevant theorem to specify RVaR and RCVaR optimization models for interval ambiguity sets on the mean returns and covariance matrix.

Definition 4 (Interval ambiguity for mean returns and covariance matrix) Mean returns and covariance matrix belong to the following interval set:

$$U_I = \{(\bar{\mu}, \bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid \bar{\mu}_- \leq \bar{\mu} \leq \bar{\mu}_+,\ \bar{\Gamma}_- \leq \bar{\Gamma} \leq \bar{\Gamma}_+\},$$

where $\bar{\mu}_-$, $\bar{\mu}_+$, $\bar{\Gamma}_-, \bar{\Gamma}_+$ are given vectors and matrices and the inequalities are component-wise. We assume there is at least one $(\bar{\mu}, \bar{\Gamma}) \in U_I$ for which $\bar{\Gamma} \succeq 0$.

Remark 9 The interval ambiguity set can be written as $U_{I_1} \times U_{I_2}$, where

$$U_{I_1} = \{\bar{\mu} \in \mathbb{R}^n \mid \bar{\mu}_- \leq \bar{\mu} \leq \bar{\mu}_+\}$$

$$U_{I_2} = \{\bar{\Gamma} \in \mathbb{S}^n \mid \bar{\Gamma}_- \leq \bar{\Gamma} \leq \bar{\Gamma}_+\}$$

are interval ambiguity sets for means and covariance matrix, respectively.

Theorem 5 *If random variable ξ has a distribution from the set $\mathbb{D}$ and $(\bar{\mu}, \bar{\Gamma}) \in U_I$. Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) for interval ambiguity are represented by the following SDP:*

$$
\begin{aligned}
\min_{\upsilon\in\mathbb{R},\, x_+,\, x_-\in\mathbb{R}^n,\, \Lambda,\, \Lambda+,\, \Lambda_-\in\mathbb{S}^n} \quad & tr(\Lambda_+\bar{\Gamma}_+) - tr(\Lambda_-\bar{\Gamma}_-) + \\
& \frac{\alpha}{1-\alpha}\upsilon + (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_- \\
s.t. \quad & \\
& \begin{bmatrix} \Lambda & \frac{x_- - x_+}{2} \\ \frac{(x_- - x_+)}{2}^\top & \upsilon \end{bmatrix} \succeq 0, \\
& \Lambda \preceq \Lambda_+ - \Lambda_-, \\
& (\bar{\mu}_- - r_f e)^\top x_- - (\bar{\mu}_+ - r_f e)^\top x_+ \geq d - r_f, \\
& x_+, x_- \geq 0,\ \Lambda,\ \Lambda_+,\ \Lambda_- \succeq 0.
\end{aligned}
\tag{22}
$$

Proof. *See Appendix A.4.*

4 Numerical tests

We illustrate the performance of the robust models and compare the robust models vis-a-vis the nominal (non-robust) models on two financial applications. First, we use simulations to test the robustness of solutions under mis-specification of mean, variance, skewness, and kurtosis of the return distributions. This test also highlights the conservativeness of the robust models and justifies the heuristic for constructing ambiguity sets. Second, we use sovereign CDS spread returns for Eurozone core and periphery, and Central, Eastern and South-Eastern Europe countries, covering the eurozone crisis period, to test the robustness of buy-and-hold and active management investment strategies obtained using the models. Since the CDS spread returns suffer regime switching before and after the crisis, they provide a natural set of data for stress-testing the robust models. We also illustrate the estimation of ellipsoids with the algorithm and the heuristic, and provide results with the execution times to understand the computational requirements of the new models. All computations were performed using MATLAB 7.14.0 on a Core $i7$ CPU 2.5GHz laptop with 8GB of RAM. SOCPs are solved using CVX and (mixed integer) linear programs with CPLEX.

4.1 Robustness under distribution ambiguity: moment mis-specification

We demonstrate the robustness of RCVaR optimal portfolios to mis-specification in the first four marginal moments. Mis-specification of higher moments is a form of distribution ambiguity and these tests illustrate robustness with respect to distribution ambiguity. A more interesting interpretation of our results is in conjunction with the work of Kaut et al. (2007), where it was established that CVaR optimization models are sensitive to mis-specification of means, covariance, skewness, and, less so, to kurtosis. Our results show that these sensitivities are eliminated by the robust models of this paper.

We consider CVaR and RCVaR optimization with a minimum expected return constraint, no short-selling, and a budget constraint. We perturb one moment at a time while keeping all other moments fixed to their original (assumed "true") value, and repeat the perturbation 100 times. Data are from Kaut et al. (2007) (see Appendix B, Table 2) for international investment portfolio, which are assumed to be the true values of the moments.

We test as follows the impact of moment mis-specification on the models:

Step 0: Fix parameter θ and define $\theta\%$ error on a moment as[4]

$$\text{true value}\,(1+\epsilon\frac{\theta}{100}),\ \epsilon \in U[-1,1].$$

Generate 100 perturbations for one moment by randomly generating ϵ, while all other moments are fixed to their true values.

Step 1: Generate 2000 scenarios using Pearson random numbers with the specified mean, standard deviation, skewness, or kurtosis for each one of the 100 perturbations from Step 0. Record the scenario sets $\{\bar{R}_k\}_{k=1}^{100}$ and their means and covariance $\{(\bar{\mu}_k, \bar{\Gamma}_m)\}_{k=1}^{100}$.

Step 2: Apply the algorithm of subsection 3.2.1 to $\{(\bar{\mu}_k, \bar{\Gamma}_k)\}_{k=1}^{100}$, to find the center $(\hat{\mu}, \hat{\Gamma})$ and the parameter δ. Chose the point with the smallest distance δ_k from the center, with its scenario set $\hat{R}$, as the *reference* scenario set.

Step 3: Solve the model on the reference scenario set and the robust counterpart, and record the optimal portfolios.

Step 4: Compute return and risk measures of the optimal portfolios over $\{\bar{R}_k\}_{k=1}^{100}$.

The model on the reference set is a proxy for the nominal model since we do not have a scenario set corresponding to the center of the ellipsoid generated by the algorithm. When using the heuristic to compute the ellipsoid we have the scenario set for the center and hence we have exactly the nominal model. The performances of the proxy and the nominal models do not differ significantly, and in the experiments we compare the robust model with the proxy.

4.1.1 Mean and variance mis-specification

Applying the simulation procedure outlined above for the first two moments is not essential, since the robustness of RVaR and RCVaR have been established theoretically. Nevertheless, we perform simulations on the first two moments to illustrate another feature of the models. Robust models are conservative and we illustrate how to control conservativeness by adjusting δ. We pick δ so that the ambiguity set is large enough to contain all 100 perturbations or only 90 or 80 and so on, by choosing in Step 2 of the algorithm the appropriate quantile of δ_k, $k = 1, \ldots, K$.

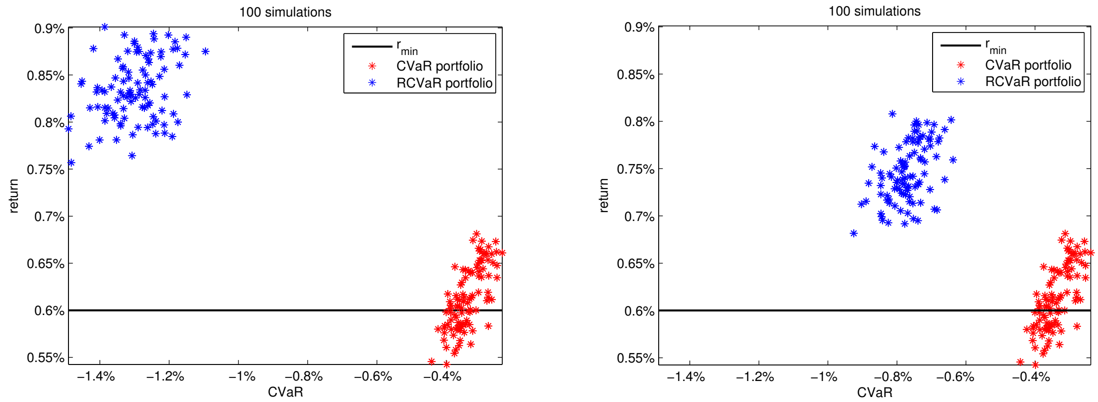
The results of simulating perturbations in the mean and variance by $\theta = 10\%$ and 20%, respectively, and different values of δ, are illustrated in Figures 2–3. We observe from panel (a) that the robust portfolio never violates the minimum return constraint out-of-sample, while the nominal portfolio violates the minimum return for about half of the perturbations. On the other hand, the robust optimal portfolio is conservative and mean returns are significantly higher than the target for all perturbations. We can remedy this situation by reducing δ using a suitable quantile of $\delta_k, k = 1, \ldots, K$. Panel (b) illustrates the effect of this modification in obtaining less conservative portfolios[5]. Running the models on the heuristic-generated ellipsoids we find them more conservative, but we do not report results as they do not provide any additional insights.

4.1.2 Skewness and kurtosis mis-specification

Kaut et al. (2007) established that CVaR portfolios are sensitive to mis-specification in the first four moments and correlations. In particular, they solve CVaR optimization models for each of 100 generated scenario sets and then evaluate all these optimal portfolios on the benchmark scenario set. Our experiment studies the same issue for RCVaR optimization. We solve the

[4] We follow Chopra and Ziemba (1993), except that they use normally distributed $\epsilon \in N[0,1]$, while we use uniformly distributed $\epsilon \in U[-1,1]$.

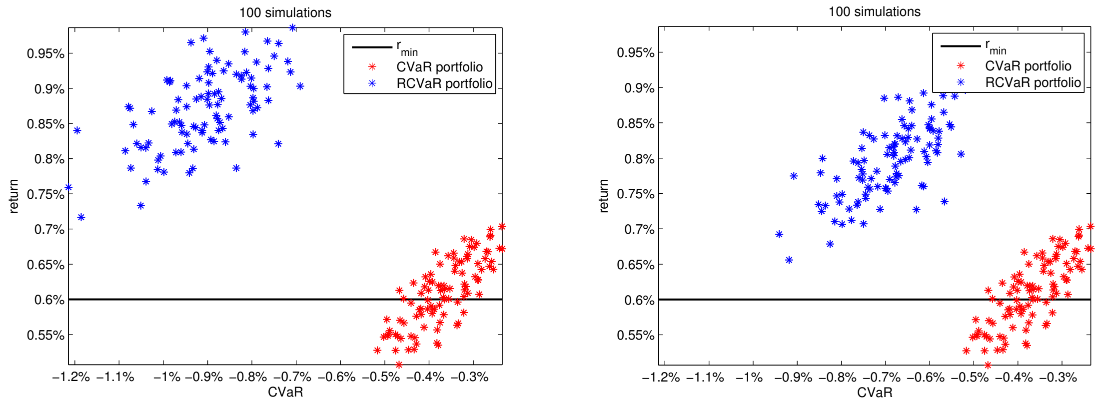
[5] Note that the less conservative portfolios still satisfy the mean return constraint, but this is now not guaranteed. What we are guaranteed is to meet the target portfolio for the chosen quantile k.



(a) $\delta = 12.24$, solution from subsection 3.2.1 algorithm

(b) $\delta = 10.61$, 80th quantile of δ_k, $k = 1, \ldots, K$

Fig. 2: Out-of-sample performance of RCVaR optimization for perturbations in means and variances with $\theta = 10\%$. Portfolios are less conservative for smaller delta.



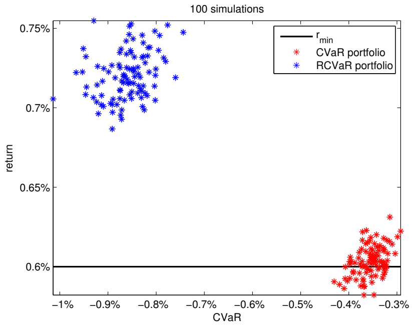
(a) $\delta = 21.13$, solution from subsection 3.2.1 algorithm

(b) $\delta = 16.94$, 80th quantile of δ_k, $k = 1, \ldots, K$

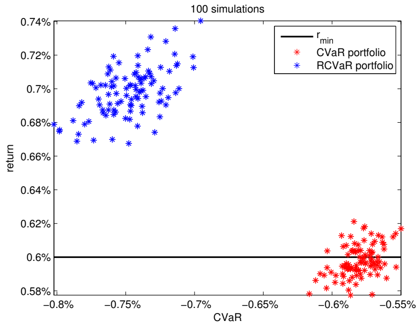
Fig. 3: Out-of-sample performance of RCVaR optimization for perturbations in means and variances with $\theta = 20\%$. Portfolios are less conservative for smaller delta.

CVaR optimization with reference scenario $\hat{R}$ and its robust counterpart, and evaluate the optimal portfolio on 100 scenario sets of skewness and kurtosis perturbations.

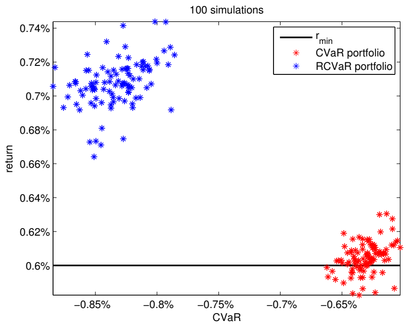
The results are illustrated in Figures 4–5. We examine the performance of optimal portfolios with increasing perturbation parameter θ, and δ chosen by the algorithm of subsection 3.2.1. Our observations are consistent with Kaut et al. (2007) on the sensitivity of CVaR strategy with respect to higher moments mis-specification, i.e., the optimal CVaR portfolio violates the minimum return constraint for the perturbed scenario sets. The sensitivities of CVaR portfolios to errors in skewness and kurtosis are in agreement with the findings of Kaut et al. (2007). The optimal RCVaR portfolios, however, satisfy the minimum return constraint even with perturbations in the higher moments. Higher moment perturbation is a form of distribution ambiguity and, hence, the results are expected since the models are robust with respect to distribution ambiguity.



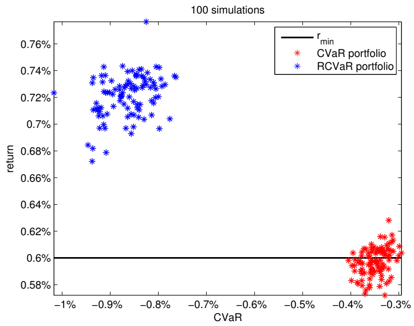
(a) $\delta = 8.49, \theta = 10\%$



(b) $\delta = 8.33, \theta = 20\%$

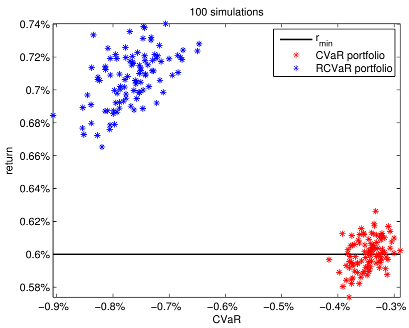


(c) $\delta = 8.45, \theta = 30\%$

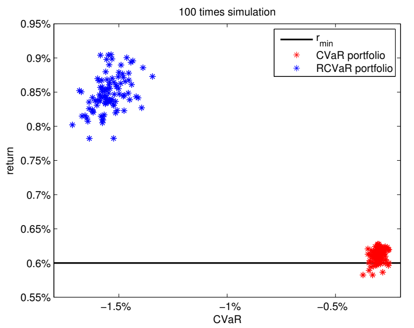


(d) $\delta = 8.42, \theta = 40\%$

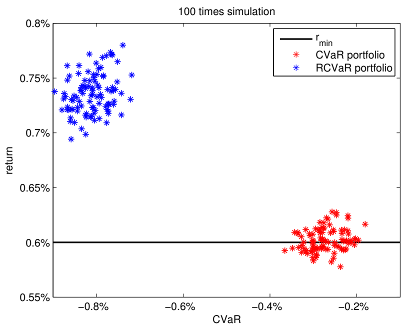
Fig. 4: Out-of-sample performance for perturbations in skewness for increasing θ. RCVaR portfolios meet the minimum return whereas CVaR portfolios do not.



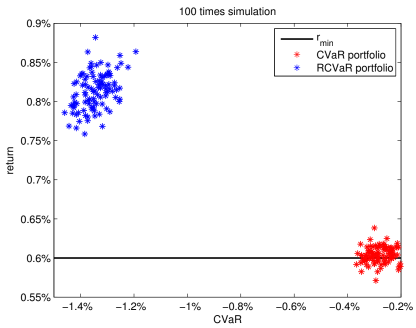
(a) $\delta = 8.55, \theta = 10\%$



(b) $\delta = 8.46, \theta = 20\%$



(c) $\delta = 8.15, \theta = 30\%$



(d) $\delta = 8.52, \theta = 40\%$

Fig. 5: Out-of-sample performance for perturbations in kurtosis for increasing θ. RCVaR portfolios meet the minimum return whereas CVaR portfolios do not.

4.2 Robustness of different investment strategies

We now construct portfolios in eight sovereign CDS for Portugal, Slovenia, Italy, Spain, Ireland, Germany, Cyprus, and Greece, using daily spread returns from 2 Feb. 2009 to 16 Sept. 2011. This period covers the eurozone crisis. Analyzing Greece CDS spreads using Bai-Perron tests we note regime switching at 20 April 2010 and 15 April 2011 (see Appendix B). Up to April 2010 there is a *tranquil* period (days 1–317), until April 2011 a *turbulent* period (days 318–575), and post April 2011 we have a *crisis* (days 576–685). This classification is convenient to stress-test the robustness of the model as the market changes from tranquil to turbulent and into a crisis, and we do so for both RVaR and RCVaR minimization without a risk free asset, no short-selling, and a budget constraint, following a buy-and-hold investment strategy.

We also use an active portfolio management strategy, to evaluate the performance of the model in real-word applications, instead of simply testing its robustness. We test the active management strategy using robust and nominal CVaR models on a set of 18 Eurozone core and periphery, and Central, Eastern and South-Eastern Europe countries. For this experiment we use the longest possible period during which there were complete data for all countries in the sample, i.e., October 2008 to March 2016. During this period again three regimes were identified, which we call *turbulent*, *crisis*, and *post-crisis*. For details on the data, identification of regime switching, and descriptive statistics during each regime see Consiglio et al. (2017). These experiments shed light to the following questions: (i) Are portfolios obtained with robust optimization models indeed robust during the crisis regime switching? (ii) How do robust portfolios perform using a well accepted risk-adjusted measure of portfolio performance (Sharpe ratio) in a realistic example, and how do they compare to their non-robust counterparts?

4.2.1 Buy-and-hold

Buy-and-hold investors use the available information to set up the model and optimize the asset allocation, which is kept fixed throughout the investment horizon. We consider an investor who develops robust models based on scenarios observed during the tranquil period and keeps them even as the markets move into turbulence and crisis. The new information is observed and used to compute portfolio performance, i.e., the risk measures VaR and CVaR, but the portfolio is not re-optimized. For each new observation we drop the oldest observation, so that the VaR and CVaR of portfolios obtained with the models are computed on a fixed-size window of the most recent data.

The ellipsoidal ambiguity set is constructed as follows. First we estimate $(\bar{\mu}, \bar{\Gamma})$ using the first 150 (out of 316) return observations in the tranquil period. Then we discard the first observation, add the 151st, and compute a new estimate of $(\bar{\mu}, \bar{\Gamma})$. This procedure is repeated by "rolling" the estimation window forward one period at a time until the end of the tranquil period. At the end of this procedure, we have 166 estimates of $(\bar{\mu}, \bar{\Gamma})$, and compute the center and δ of the ambiguity set using the algorithm of subsection 3.2.1. Nominal models use observed data over the tranquil period to obtain minimum VaR and CVaR portfolios, which are held throughout the turbulent period. We repeat this process using scenarios from the turbulent period and evaluate out-of-sample performance into the crisis period.

Results are shown in Figures 6–7. We observe that the out-of-sample risk measure for the non-robust portfolios may be larger than the in-sample value, but not so for the robust portfolios. The robust portfolios remain robust even when there is a regime switch from tranquil to turbulent and from turbulent to crisis.

The well–known conservativeness (Ben-Tal and Nemirovski, 2002; Bertsimas and Sim, 2004) of robust portfolios in this example is observed in the gap between the in-sample and out-of-sample values. The gap is unavoidable, since we consider a robust model for the worst case under ambiguity in all problem data. However, we point out that the magnitude of the gap depends on data ambiguity and model characteristics. Since we assume ambiguity in means, variances and covariances, as well as the distribution, the gap in these examples is large. Research in robust optimization typically finds small gaps. However, (Thiele, 2010, Example 1.1) stress test numerically a portfolio allocation problem and found that when more than 41 out of the 150 assets had ambiguous data, the solution would become extremely conservative with the portfolio invested in a single asset. In our application we assume that all assets have ambiguous data. Using smaller values of δ from the algorithm (or the heuristic), we get less conservative solutions, as illustrated in Figures 6–7. The gap is reduced as we decrease δ. For $\delta = 0$ and without distribution ambiguity we get the same in-sample solution as the nominal model. However, now the solution is not robust. We also note that the gap has a large component due to distribution ambiguity, manifested when we set $\delta = 0$ without eliminating distribution. It appears, in this particular example, that distribution ambiguity is an extreme assumption and models with ambiguity in means, variances and correlations, but no distribution ambiguity, would be less conservative. In financial applications it may be possible to avoid distribution ambiguity, especially in non-crisis situations, but if decision makers are concerned about a potential crisis the more general ambiguity incorporated in our models is more appropriate.

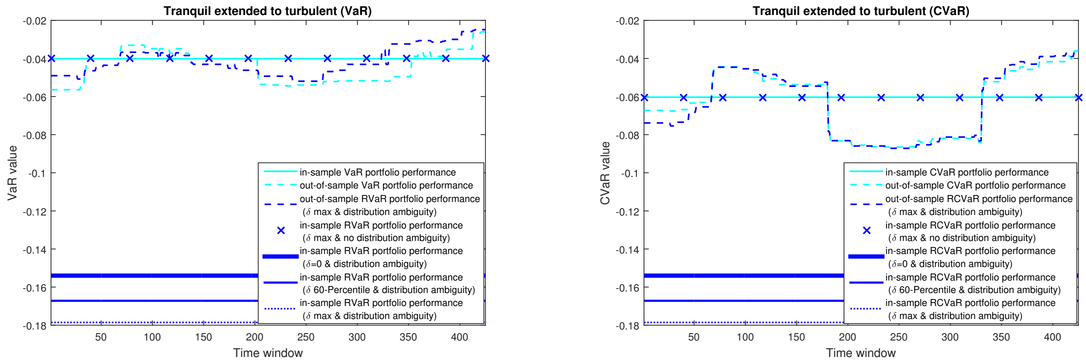
We note from this experiment that the out-of-sample performance does not differ significantly between the robust and non-robust model portfolios. (We do not report in the figures the out-of-sample performance of portfolios obtained with smaller values of δ, as they lie between the lines shown for the robust and the nominal models.) Hence, it may not be possible to gain robustness and at the same time improve portfolio performance. From the figure it appears that the robust models have slightly better performance than the non-robust counterparts out-of-sample, but this is serendipitous. To evaluate further this issue we use the models for active portfolio management in the next section.

4.2.2 Active management

Active portfolio managers use the available information to set up the model and optimize asset allocation for one time period, but as new information arrives the model data are updated and the portfolio is re-optimized. We consider an investor who starts calibrating nominal and robust models with the scenarios from the first 150 observations of the tranquil period, with $\delta = 0$ for the starting robust model. Subsequently, a new data point is observed, the oldest observation is dropped, and we compute the risk measures with the shifted window, and the portfolio ex post return for the new data point. After the time window is shifted we re-optimize the asset allocation with the new information. For the robust model we use the new information to update δ and construct an ellipsoid using the algorithm of subsection 3.2.1. This procedure is repeated until the end of the turbulent period. The same experiment is carried out starting with the first 150 observations of the turbulent period and finishing at the end of the crisis.

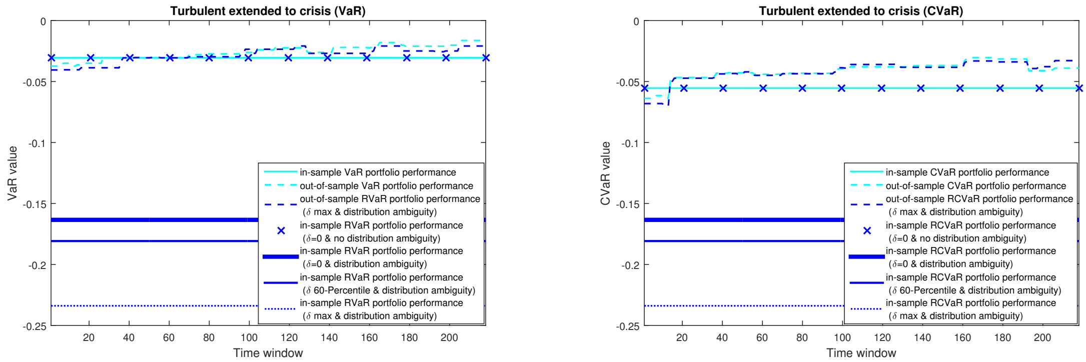
Results with the sample of eight countries are shown in Figures 8–9. Panels (a)–(b) show differences between in- and out-of-sample risk measures. The investor is on the safe side when the difference is positive and suffers unexpected losses for negative differences. Figure 8(a)-(b) shows out-of-sample performance occasionally deviating from the in-sample estimate. As the time window rolls forward the robust model registers few and minor downside violations, as a result of enlarged ambiguity sets with increasing δ (Figure 12). This improvement is less pronounced in Figure 9(a)-(b), since spreads change substantially during the crisis and learning is insufficient to build an ellipsoid containing crisis movements. Robust models can not be better than the data defining the ambiguity sets.

Figures 8(c) and 9(c) plot the ex-post cumulative growth of investments using both the robust and non-robust models. We calculate the Sharpe ratios for the returns of portfolios developed using VaR, CVaR, and their robust counterpart. We take the German 3-month treasury bill rate as the risk free in Sharpe ratio calculations, and the results are reported in the figure. We test the hypothesis that the Sharpe ratios are identical between the robust and non-robust strategies using the test established in (Wright et al., 2014), and can not reject it at the 0.01



(a) VaR and RVaR (b) CVaR and RCVaR

Fig. 6: Out-of-sample performance of buy-and-hold for tranquil-to-turbulent. ($\delta = 10.28$)

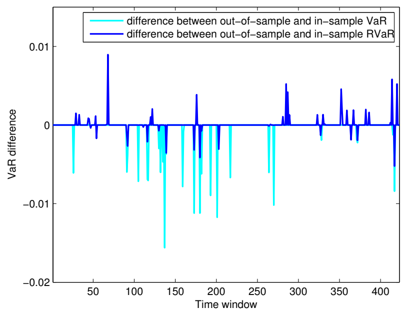


(a) VaR and RVaR (b) CVaR and RCVaR

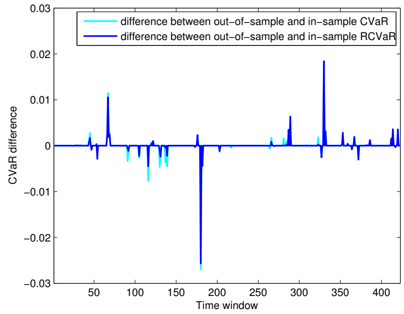
Fig. 7: Out-of-sample performance of buy-and-hold for turbulent-to-crisis. ($\delta = 21.63$)

significance level, for both tranquil-to-turbulent and turbulent-to-crisis periods (Appendix C). Robust solutions do not necessarily sacrifice portfolio performance.

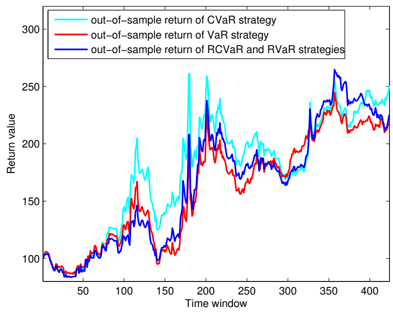
It is an unexpected positive finding that we gain robustness without sacrificing performance, and we repeat the experiment using the sample of 18 countries over the longer time period, using a regime identification that is suitable for this combination of countries and dates (turbulent, crisis, and post-crisis). Results are reported in Figure 10 and the test statistics in Appendix C, and confirm the finding. We also use a variation of Sharpe ratio that considers only downside risk (Ziemba, 2005), and find that the robust models perform better than the nominal models during the turbulent to crisis period, with one-sided ratios of 0.0048 for RCVaR and -0.0065 for CVaR (although we are not aware of statistical tests for the down-sided Sharpe ratio). These are very encouraging results for the use of robust optimization for financial applications. We hasten to emphasize, however, that this is an empirical finding on two data-sets. While it is encouraging for the use of robust models, this may not be the case for different data sets and different time periods. Indeed, one should be ready to sacrifice portfolio performance to gain portfolio robustness.



(a) VaR and RVaR

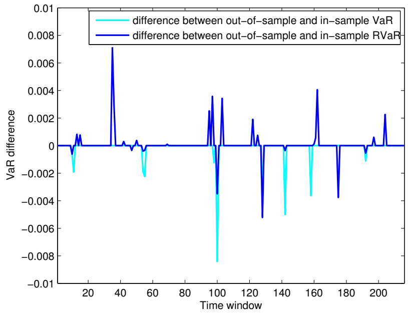


(b) CVaR and RCVaR

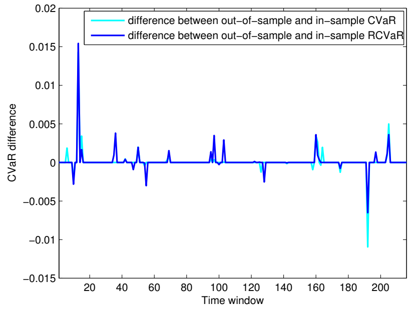


(c) Cumulative returns with Sharpe ratios: VaR=-0.093, CVaR=-0.079, RVaR=RCVaR=-0.087. The hypothesis that Sharpe ratios are identical can not be rejected at the 0.99 level.

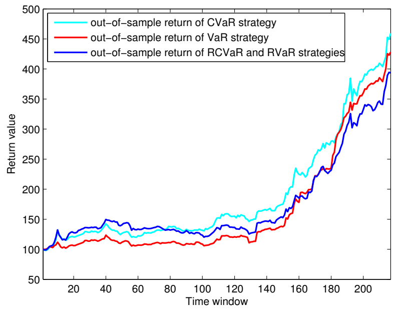
Fig. 8: Out-of-sample performance with active management for the eight-country sample during tranquil-to-turbulent.



(a) VaR and RVaR

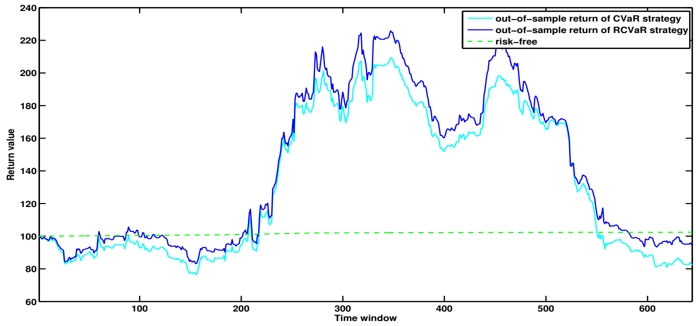


(b) CVaR and RCVaR

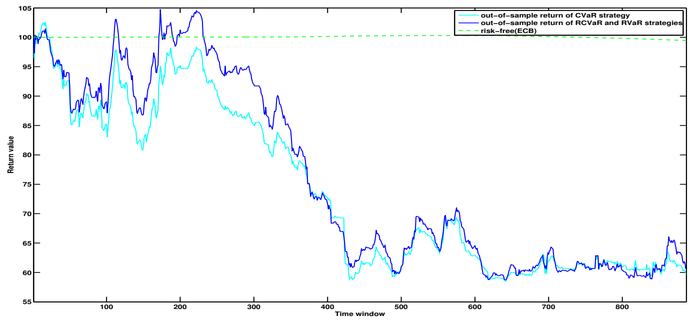


(c) Cumulative returns with Sharpe ratios: VaR=-0.127, CVaR=-0.104, RVaR=RCVaR=-0.153. The hypothesis that Sharpe ratios are identical can not be rejected at the 0.95 level.

Fig. 9: Out-of-sample performance with active management for the eight-country sample during turbulent-to-crisis.

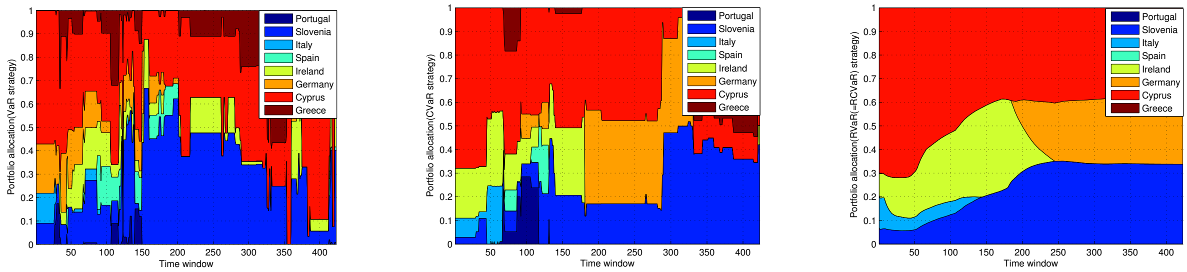


(a) Cumulative returns with Sharpe ratios: CVaR=-0.006, RCVaR=0.004. The hypothesis that Sharpe ratios are identical can not be rejected at the 0.95 level.



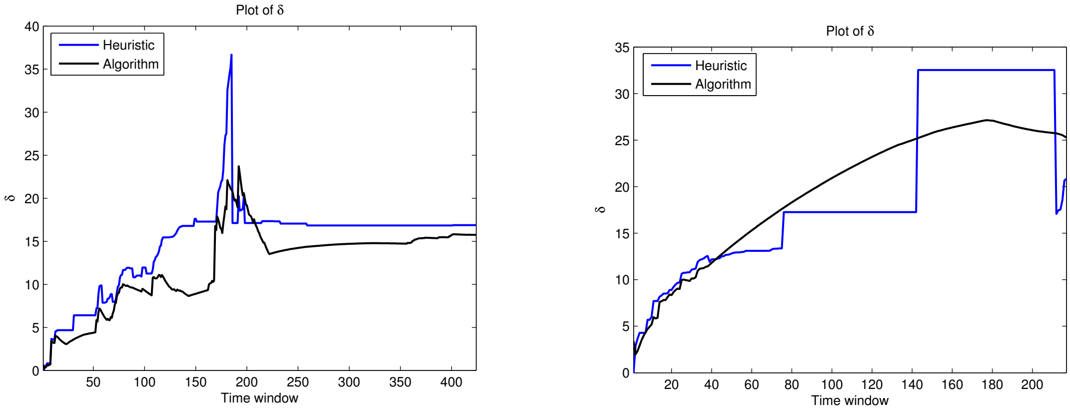
(b) Cumulative returns with Sharpe ratios: CVaR=-0.061, RCVaR=-0.060. The hypothesis that Sharpe ratios are identical can not be rejected at the 0.95 level.

Fig. 10: Out-of-sample performance with active management for the 18-country sample during (a) turbulent-to-crisis, and (b) crisis-to-post-crisis.



(a) VaR (b) CVaR (c) RVaR=RCVaR

Fig. 11: Portfolio composition with different models for tranquil-to-turbulent. Portfolio turnover for VaR=0.09, CVaR=0.03 and RVaR=RCVaR=0.004.



(a) Tranquil-to-turbulent (b) Turbulent-to-crisis

Fig. 12: Values of δ estimated with the algorithm and the heuristic.

4.2.3 Stability of optimal portfolios

We use the active management experiment to revisit the instability issue discussed in Section 2. Figure 11 illustrates the portfolio composition in the eight-country sample for the tranquil-to-turbulent period. The robust portfolios change gradually, but not so the nominal counterparts. Portfolio turnover of the robust model is 0.004, an order of magnitude smaller than that of VaR (0.09) or CVaR (0.03). The improved stability of the portfolios generated by the robust models was noticeable in all experiments.

4.3 Estimating the ellipsoids

We report in Figure 12 the values of δ obtained with the algorithm and the heuristic during the active management experiments on the eight-country sample. The algorithm generates tighter ellipsoids for tranquil-to-turbulent period, while there is no clear advantage of one method over the other in turbulent-to-crisis period. In all experiments reported above we use ellipsoids constructed by the algorithm[6].

[6] We also performed experiments using heuristic-constructed ellipsoids, without any significant differences.

4.4 Computational requirements

The robust CVaR optimization models are computationally tractable. By solving robust CVaR optimization models we get robust solutions for CVaR (and VaR) optimization, with marginally larger computation time than solving a nominal CVaR model. The RCVaR model, which is identical to RVaR, requires significantly less computational resources than the nominal VaR model which is a mixed integer program. Table 1 summarizes model sizes and solution times for all models.

Problem	Model	Number of constraints	Number of variables	Number of binary variables
8-country problem with 100 scenarios				
VaR	MIP	152	310	150
CVaR	LP	151	309	0
RCVaR=RVaR	SOCP	1	8	0
18-country problem with 500 scenarios				
VaR	MIP	502	1020	500
CVaR	LP	501	1019	0
RCVaR $\doteq$ RVaR	SOCP	1	18	0

Model and solver	8-country model (seconds)	18-country model (seconds)
CVaR using CPLEX	1.96	4.31
VaR using CPLEX	33.25	206.94
RCVaR=RVaR using CVX	51.40	56.20

Table 1: Model dimensions (top) and solution times (bottom) for solving 100 instances of the models.

5 Conclusions

This paper develops models for robust optimization of VaR and CVaR for the most general ambiguity sets, namely joint ambiguity in the distribution, mean returns, and covariance matrix. RVaR and RCVaR optimization for distribution ambiguity reduce to the same second order cone program. This result allows us to develop several tractable models using ellipsoidal, polytopic, and interval ambiguity sets for mean returns and covariance matrix. These models expand the arsenal of robust optimization tools for risk management.

The paper also suggests an algorithm and a heuristic to construct joint ellipsoidal ambiguity sets from a set of point estimates. We show how to control the size of the ellipsoid, thus limiting the well known conservativeness of robust optimization models.

Numerical results support the following conclusions:

1. RCVaR optimal portfolios are robust with respect to mis-specifications in the first four moments.
2. RVaR and RCVaR models could produce conservative solutions under the extreme conditions assumed in the model, namely ambiguity in all moments and the underlying distributions. If users can resolve some of this ambiguity then less general models are more appropriate. In any event, both the algorithm and the heuristic for constructing an ellipsoidal ambiguity set provide a way to select the ellipsoidal parameter in order to control conservativeness.
3. Buy-and-hold investment strategies based on robust optimization models perform well even out-of-sample and under the extreme market movements of a financial crisis.
4. Active investment strategies based on robust models do not have inferior performance to strategies based on nominal models, but they are less sensitive to out-of-sample perturbances. Hence, we gain robustness without sacrificing portfolio performance , even under the extreme conditions assumed by the model, and this is a very positive finding.

The last finding is very encouraging but is based on limited experimentation and deserves further empirical investigation. It also begs for an answer to the question whether robust portfolios always have non-inferior performance to their nominal counterparts in efficient markets.

References

P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

A. Ben-Tal and A. Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization.* Princeton University Press, Princeton, NJ, 2009.

D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

D. Bertsimas, G. J. Lauprete, and A. Samarov. Shortfall as a risk measure: Properties, optimization and applications. *Journal of Economic Dynamics and Control*, 28(7):1353–1381, 2004.

D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

M. J. Best and R. R. Grauer. Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8):980–989, 1991.

J. Čerbáková. Worst-case VaR and CVaR. In *Operations Research Proceedings 2005*, pages 817–822. Springer Berlin Heidelberg, 2006.

S. Ceria and R. A. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2):109–127, 2006.

L. Chen, S. He, and S. Zhang. Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research*, 59(4):847–865, 2011.

V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11, 1993.

A. Consiglio, S. Lotfi, and S. Zenios. Portfolio diversification in the sovereign CDS market. *Annals of Operations Research*, (Published on line July 27), 2017.

E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.

L. El Ghaoui, M. Oks, and F. Oustry. Worst-case Value-at-Risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

F. J. Fabozzi, D. Huang, and G. Zhou. Robust portfolios: contributions from operations research and finance. *Annals of Operations Research*, 176(1):191–220, 2010.

D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

J.-Y. Gotoh, K. Shinozaki, and A. Takeda. Robust portfolio techniques for mitigating the fragility of CVaR minimization and generalization to coherent risk measures. *Quantitative Finance*, 13(10):1621–1635, 2013.

M. Kaut, H. Vladimirou, S. W. Wallace, and S. A. Zenios. Stability analysis of portfolio management with conditional value-at-risk. *Quantitative Finance*, 7(4):397–409, 2007.

F. H. Knight. *Risk, Uncertainty and Profit.* Houghton Mifflin Company, Boston, MA, 1921.

S. Kou, X. Peng, and C. C. Heyde. External risk measures and Basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.

A. Lim, J. Shanthikumar, and G. Y. Yan. Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163–171, 2011.

Z. Lu. A computational study on robust portfolio selection based on a joint ellipsoidal uncertainty set. *Mathematical Programming*, 126(1):193–201, 2011.

J. S. Marron and M. P. Wand. Exact mean integrated squared error. *The Annals of Statistics*, 20(2):712–736, 1992.

J. Mulvey, R. Vanderbei, and S. Zenios. Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281, 1995.

A. B. Paç and M. Ç. Pınar. Robust portfolio choice with CVaR and VaR under distribution and mean return ambiguity. *TOP*, 22(3):875–891, 2014.

G. C. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In S. Uryasev, editor, *Probabilistic Constrained Optimization*, pages 272–281. Kluwer Academic Publishers, 2000.

A. G. Quaranta and A. Zaffaroni. Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance*, 32(10):2046–2056, 2008.

R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 2002.

K. Schottle and R. Werner. Robustness properties of mean-variance portfolios. *Optimization*, 58(6):641–663, 2009.

A. Thiele. A note on issues of over-conservatism in robust optimization with cost uncertainty. *Optimization*, 59(7):1033–1040, 2010.

R. H. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132 (1-4):157–187, 2004.

J. A. Wright, S.-C. P. Yam, and S. Yung. A test for the equality of multiple Sharpe ratios. *Journal of Risk*, 16(4):3–21, 2014.

K. Ye, P. Parpas, and B. Rustem. Robust portfolio optimization: a conic programming approach. *Computational Optimization and Applications*, 52(2):463–481, 2012.

S. Zenios. *Practical Financial Optimization. Decision making for financial engineers.* Blackwell-Wiley Finance, Malden, MA, 2007.

L. Zhu, Y. Li, and T. F. Coleman. Min-max robust and CVaR robust mean-variance portfolios. *Journal of Risk*, 11(3), 2008.

S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168, 2009.

W. T. Ziemba. The symmetric downside risk Sharpe ratio. *Journal of Portfolio Management*, 32(1):108–122, 2005.

S. Zymler, D. Kuhn, and B. Rustem. Worst-case value at risk of nonlinear portfolios. *Management Science*, 59(1):172–188, 2013a.

S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013b.

A Appendix: Proofs

A.1 Proof of Theorem 2

To formulate the robust counterpart of (5) and (6) we need an explicit formulation of (10) and (11), respectively. Using Proposition 1 we write both RVaR and RCVaR models (10) and (11) as

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \max_{(\bar\mu,\bar\Gamma)\in U_\delta(\hat\mu,\hat\Gamma)} \quad & -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x} \\ s.t. \quad & \min_{(\bar\mu,\bar\Gamma)\in U_\delta(\hat\mu,\hat\Gamma),\pi\in\mathbb{D}} (\bar\mu-r_fe)^\top x\ge d-r_f. \end{aligned} \tag{23}$$

To find an explicit formulation we need the optimal value of the inner problem

$$\max_{(\bar\mu,\bar\Gamma)\in U_\delta(\hat\mu,\hat\Gamma)} -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}. \tag{24}$$

We decompose the robustification first in $\bar\mu$ and then in $\bar\Gamma$ via an additional parameter $\kappa\in(0,1)$, where $U_{\sqrt{\kappa}\delta}(\hat\mu)$, $U_{\sqrt{1-\kappa}\delta}(\hat\Gamma)$ are defined as in Remark 3. It is easy to see that (24) is equivalent to:

$$\max_{\kappa\in[0,1]}\max_{\bar\Gamma\in U_{\sqrt{1-\kappa}\delta}(\hat\Gamma)}\max_{\bar\mu\in U_{\sqrt{\kappa}\delta}(\hat\mu)} -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x}. \tag{25}$$

We start with the innermost maximization problem

$$\max_{\bar\mu\in U_{\sqrt{\kappa}\delta}(\hat\mu)} -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x},$$

or, equivalently,

$$\begin{aligned} \max\nolimits_{\bar\mu\in\mathbb{R}^n} \; & -r_f-(\bar\mu-r_fe)^\top x+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x} \\ \text{s.t.} \quad & S(\bar\mu-\hat\mu)^\top\hat\Gamma^{-1}(\bar\mu-\hat\mu)\le\kappa\delta^2. \end{aligned} \tag{26}$$

This is maximized for $\bar\mu^*=\hat\mu-\delta\sqrt{\frac{\kappa}{S}}\frac{\hat\Gamma^{\frac12}x}{\|\hat\Gamma^{\frac12}x\|}$. Plugging in this solution we obtain the middle maximization problem as

$$\begin{aligned} \max_{\bar\Gamma\in\mathbb{S}^n} \; & -r_f-(\hat\mu-r_fe)^\top x+\delta\sqrt{\frac{\kappa}{S}}\|\hat\Gamma^{\frac12}x\|+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\bar\Gamma x} \\ s.t. \quad & \frac{S-1}{2}\|\hat\Gamma^{-\frac12}(\bar\Gamma-\hat\Gamma)\hat\Gamma^{-\frac12}\|_{tr}^2\le(1-\kappa)\delta^2. \end{aligned}$$

Using the variable transformation $\bar{\bar\Gamma}=\hat\Gamma^{-\frac12}(\bar\Gamma-\hat\Gamma)\hat\Gamma^{-\frac12}$, this becomes

$$\begin{aligned} \max_{\bar{\bar\Gamma}\in\mathbb{S}^n} \; & -r_f-(\hat\mu-r_fe)^\top x+\delta\sqrt{\frac{\kappa}{S}}\|\hat\Gamma^{\frac12}x\|+\frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top\hat\Gamma x+x^\top\hat\Gamma^{\frac12}\bar{\bar\Gamma}\hat\Gamma^{\frac12}x} \\ \text{s.t.} \; & \|\bar{\bar\Gamma}\|_{tr}^2\le\frac{2}{S-1}(1-\kappa)\delta^2. \end{aligned} \tag{27}$$

Since the square root function is monotonically increasing, the objective function of (27) is maximized if and only if $x^\top\hat\Gamma^{\frac12}\bar{\bar\Gamma}\hat\Gamma^{\frac12}x$ is maximized. Thus, we let $y:=\hat\Gamma^{\frac12}x$ and solve

$$\max\nolimits_{\bar{\bar\Gamma}\in\mathbb{S}^n} \; y^\top\bar{\bar\Gamma}y \tag{28}$$

$$s.t. \quad \|\bar{\bar\Gamma}\|_{tr}^2\le\frac{2}{S-1}(1-\kappa)\delta^2. \tag{29}$$

The optimal solution $\bar{\bar\Gamma}^*$ is given by

$$\bar{\bar\Gamma}^*=\delta\sqrt{\frac{2}{S-1}(1-\kappa)}\frac{y}{\|y\|}.\frac{y^\top}{\|y\|}.$$

Plugging everything back we get the optimal value of (27) as

$$-r_f - (\hat{\mu} - r_f e)^\top x + \delta\sqrt{\frac{\kappa}{S}}\|\hat{\Gamma}^{\frac{1}{2}}x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{(1+\delta\sqrt{\frac{2(1-\kappa)}{S-1}}}\|\hat{\Gamma}^{\frac{1}{2}}x\|,$$

which is substituted back in problem (25) to get

$$\begin{aligned}
&\max_{\kappa\in[0,1]} -r_f - (\hat{\mu} - r_f e)^\top x + \delta\sqrt{\frac{\kappa}{S}}\|\hat{\Gamma}^{\frac{1}{2}}x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{(1+\delta\sqrt{\frac{2(1-\kappa)}{S-1}}}\|\hat{\Gamma}^{\frac{1}{2}}x\| \\
&= -r_f - (\hat{\mu} - r_f e)^\top x + \max_{\kappa\in[0,1]}\left(\delta\sqrt{\frac{\kappa}{S}} + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{(1+\delta\sqrt{\frac{2(1-\kappa)}{S-1}}}\right)\|\hat{\Gamma}^{\frac{1}{2}}x\| \\
&= -r_f - (\hat{\mu} - r_f e)^\top x + \left(\max_{\kappa\in[0,1]} f(\kappa)\right)\|\hat{\Gamma}^{\frac{1}{2}}x\|,
\end{aligned}$$

which is the objective function of (15).

Now, the robust counterpart of minimum return constraint is equivalent to:

$$\min_{\kappa\in[0,1]}\ \min_{\bar{\mu}\in U_{\sqrt{\kappa}\delta}(\hat{\mu})}(\bar{\mu} - r_f e)^\top x \geq d - r_f.$$

Following the same course as in solving the inner problem (24), we get:

$$\min_{\kappa\in[0,1]}(\hat{\mu} - r_f e)^\top x - \delta\sqrt{\frac{\kappa}{S}}\|\hat{\Gamma}^{\frac{1}{2}}x\| \geq d - r_f.$$

The solution is $\kappa = 1$ and the robust counterpart of minimum return constraint is:

$$(\hat{\mu} - r_f e)^\top x - \frac{\delta}{\sqrt{S}}\|\hat{\Gamma}^{\frac{1}{2}}x\| \geq d - r_f,$$

which is the constraint in (15). This completes the proof.

A.2 Proof of Theorem 3

First we state some well-known properties of Kronecker product $\otimes$.

Proposition 2 *Assume A, B, C, D and X are given matrices of conformable sizes.*

(i) $tr(AB) = tr(BA)$

(ii) $tr(A^\top B) = vec(A)^\top vec(B)$

(iii) $vec(AXB) = (B^\top \otimes A)vec(X)$

(iv) $(B \otimes A)(C \otimes D) = BC \otimes AD$

(v) $(B \otimes A)(C \otimes D) = BC \otimes AD$

where $vec(A)$ denotes the vector obtained by stacking the columns of $A \in \mathbb{R}^{m\times n}$ successively underneath each other.

We transform the problem using new variables $(\hat{\mu}, \Gamma^-)$ where $\Gamma^- = \hat{\Gamma}^{-1}$ is also positive definite, and develop the analysis on the transformed equivalent problem. We show that the transformed problem is convex and Slater condition hold. Therefore, we use the KKT optimality conditions to derive the optimal solution of the transformed problem and compute the inverse of Γ^- to obtain $\hat{\Gamma}$.

Let $H = H(\hat{\mu}, \hat{\Gamma})$ denote the objective function in (17). Then

$$H = \sum_{k=1}^{K} S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2}tr(\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-1}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}). \tag{30}$$

Using property (i) we get

$$\begin{aligned} H &= \sum_{k=1}^{K} S\, tr(\hat{\Gamma}^{-1}(\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top) + \frac{S-1}{2} tr(\hat{\Gamma}^{-1}(\bar{\Gamma}_k-\hat{\Gamma})\hat{\Gamma}^{-1}(\bar{\Gamma}_k-\hat{\Gamma})) \\ &= \sum_{k=1}^{K} S\, tr(\hat{\Gamma}^{-1}(\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top) + \frac{S-1}{2} tr((\hat{\Gamma}^{-1}\bar{\Gamma}_k - I)^2). \end{aligned}$$

Applying properties (ii)–(iii) we get

$$\begin{aligned} H &= \sum_{k=1}^{K} S\, vec((\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top)^\top vec(\hat{\Gamma}^{-1}) + \frac{S-1}{2} vec(\bar{\Gamma}_k\hat{\Gamma}^{-1}-I)^\top vec(\hat{\Gamma}^{-1}\bar{\Gamma}_k - I) \\ &= \sum_{k=1}^{K} S\, vec((\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top)^\top vec(\hat{\Gamma}^{-1}) + \\ &\quad \sum_{k=1}^{K} \frac{S-1}{2}((I\otimes\bar{\Gamma}_k)vec(\hat{\Gamma}^{-1}) - vec(I))^\top((\bar{\Gamma}_k\otimes I)vec(\hat{\Gamma}^{-1}) - vec(I)). \end{aligned}$$

Finally, by replacing $\hat{\Gamma}^{-1}$ by Γ^{-}, doing some straightforward calculations, and using properties (iv)–(v) we get the new formulation of $H(\hat{\mu},\hat{\Gamma})$ in terms of $(\hat{\mu},\Gamma^{-})$, which we call $G(\hat{\mu},\Gamma^{-})$,

$$\begin{aligned} G(\hat{\mu},\Gamma^{-}) = vec(\Gamma^{-})^\top &\left[\frac{S-1}{2}\sum_{k=1}^{K}\bar{\Gamma}_k\otimes\bar{\Gamma}_k\right] vec(\Gamma^{-}) + \frac{nK(S-1)}{2} + \\ &\left[S\sum_{k=1}^{K} vec((\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top) - (S-1)\sum_{k=1}^{K} vec(\bar{\Gamma}_k)\right]^\top vec(\Gamma^{-}). \end{aligned} \tag{31}$$

Now that we have H as a function of $(\hat{\mu},\Gamma^{-})$, we write the transformed problem as

$$\min_{\hat{\mu}\in\mathbb{R}^n,\, \Gamma^{-}\in\mathbb{S}^n_{++}} G(\hat{\mu},\Gamma^{-}).$$

One can easily check that the Hessian matrix of function G is positive semidefinite and Slater condition holds, hence KKT conditions give us the optimal solution. There are no constraints on mean returns, and Γ^{-}, being the inverse of a positive definite matrix, is positive definite and in the interior of the positive semidefinite cone. Hence, the KKT optimality conditions reduce to:

$$\begin{aligned} \nabla G_{\hat{\mu}} &= 0, \\ \nabla G_{vec(\Gamma^{-})} &= 0. \end{aligned} \tag{32}$$

To obtain $\nabla G_{\hat{\mu}}$ we take the differential with respect to $\hat{\mu}$:

$$\begin{aligned} dG &= d\sum_{k=1}^{K} S\, vec((\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top)^\top vec(\Gamma^{-}) \\ &= \sum_{k=1}^{K} S\, vec(d((\hat{\mu}-\bar{\mu}_k)(\hat{\mu}-\bar{\mu}_k)^\top))^\top vec(\Gamma^{-}) \\ &= \sum_{k=1}^{K} S\, \left(vec(d\hat{\mu}(\hat{\mu}-\bar{\mu}_k)^\top)^\top vec(\Gamma^{-}) + vec((\hat{\mu}-\bar{\mu}_k)(d\hat{\mu})^\top)^\top vec(\Gamma^{-})\right). \end{aligned}$$

Using properties (ii)–(iii) we get:

$$\begin{aligned} dG &= \sum_{k=1}^{K} S\, \left(tr((\hat{\mu}-\bar{\mu}_k)(d\hat{\mu})^\top\Gamma^{-}) + tr(d\hat{\mu}(\hat{\mu}-\bar{\mu}_k)^\top\Gamma^{-})\right) \\ &= \sum_{k=1}^{K} S\, \left(tr((\hat{\mu}-\bar{\mu}_k)(d\hat{\mu})^\top\Gamma^{-}) + tr(d\hat{\mu}(\hat{\mu}-\bar{\mu}_k)^\top\Gamma^{-})\right) \\ &= \sum_{k=1}^{K} 2S\,(\hat{\mu}-\bar{\mu}_k)^\top\Gamma^{-} d\hat{\mu}. \end{aligned}$$

Hence,

$$\nabla G_{\hat{\mu}} = (\frac{dG}{d\hat{\mu}})^\top = \sum_{k=1}^{K} 2S\,\Gamma^{-}(\hat{\mu} - \bar{\mu}_k).$$

Therefore the solution of $\nabla G_{\hat{\mu}} = 0$ is $\hat{\mu} = \frac{1}{K}\sum_{k=1}^{K}\bar{\mu}$.

Calculating $\nabla G_{vec(\Gamma^{-})}$ we obtain the second equation in (32) as:

$$\left[(S-1)\sum_{k=1}^{K}\bar{\Gamma}_k \otimes \bar{\Gamma}_k\right] vec(\Gamma^{-}) + \left[S\sum_{k=1}^{K} vec((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) - (S-1)\sum_{k=1}^{K} vec(\bar{\Gamma}_k)\right] = 0,$$

which suggests the following system of linear equations in terms of $vec(\Gamma^{-})$:

$$\left[\sum_{k=1}^{K}\bar{\Gamma}_k \otimes \bar{\Gamma}_k\right] vec(\Gamma^{-}) = \left[\sum_{k=1}^{K} vec(\bar{\Gamma}_k) - \frac{S}{(S-1)}\sum_{k=1}^{K} vec((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top)\right] = 0.$$

All $\bar{\Gamma}_k,\ k \in \{1,\ldots,K\}$ are positive semidefinite matrices. The sum and Kronecker product of two positive semidefinite matrices are positive semidefinite matrices, thus $\sum_{k=1}^{K}\bar{\Gamma}_k \otimes \bar{\Gamma}_k$ is positive semidefinite. To prove uniqueness of solution, assume $\bar{\Gamma}_l$, for some $l \in \{1,\ldots,K\}$ is a positive definite matrix, then so is $\bar{\Gamma}_l \otimes \bar{\Gamma}_l$. Also, $\bar{\Gamma}_k \otimes \bar{\Gamma}_k$, for all $k \in \{1,\ldots,K\}/\{l\}$ are positive semidefinite matrices. These all together imply that $\sum_{k=1}^{K}\bar{\Gamma}_k \otimes \bar{\Gamma}_k$ is a positive definite matrix, that is the coefficient matrix is a full rank matrix and thus (18) has a unique solution.

A.3 Proof of Theorem 4

By Proposition (1), RVaR and RCVaR optimization models can be written as:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \max_{(\bar{\mu},\bar{\Gamma})\in U_P} \quad & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top \bar{\Gamma} x} \\ s.t. \quad & \min_{(\bar{\mu},\bar{\Gamma})\in U_P} r_f + (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \tag{33}$$

Using the representation $U_{P_1} \times U_{P_2}$ of U_P, one can easily see that the inner optimization problems appeared in objective function and constraint can be decomposed into easier subproblems and thus we get:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} -r_f - \min_{\bar{\mu}\in U_{P_1}} \quad & (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \max_{\bar{\Gamma}\in U_{P_2}} \sqrt{x^\top \bar{\Gamma} x} \\ s.t. \quad & r_f + \min_{\bar{\mu}\in U_{P_1}} (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \tag{34}$$

Obviously, $\min_{\bar{\mu}\in U_{P_1}} (\bar{\mu} - r_f e)^\top x = \min_{1\leq j\leq J} (\bar{\mu}_j - r_f e)^\top x$ and

$$\max_{\bar{\Gamma}\in U_{P_2}} \sqrt{x^\top \bar{\Gamma} x} = \max_{1\leq j\leq J}\sqrt{x^\top \bar{\Gamma}_j x}.$$

Letting $\beta = r_f + \min_{1\leq j\leq J} (\bar{\mu}_j - r_f e)^\top x$ and $\omega = \max_{1\leq j\leq J}\sqrt{x^\top \bar{\Gamma}_j x}$, we get the result.

A.4 Proof of Theorem 5

By Proposition (1), RVaR and RCVaR optimization models can be written as:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \max_{(\bar{\mu},\bar{\Gamma})\in U_I} \quad & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}\sqrt{x^\top \bar{\Gamma} x} \\ s.t. \quad & \min_{(\bar{\mu},\bar{\Gamma})\in U_I} r_f + (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \tag{35}$$

We start with the objective function. Using Theorem 1 of El Ghaoui et al. (2003) we replace the inner maximization objective function by the following:

$$\begin{aligned} \min_{\upsilon\in\mathbb{R},\,\Lambda\in\mathbb{S}^n_+} \quad & tr(\Lambda\bar{\Gamma}) + \frac{\alpha}{1-\alpha}\upsilon - (r_f + (\bar{\mu} - r_f e)^\top x) \\ s.t. \quad & \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & \upsilon \end{bmatrix} \succeq 0, \end{aligned} \tag{36}$$

Hence, the inner maximization in the objective function of model (35) is equivalent to:

$$\begin{aligned} \max_{(\bar{\mu},\bar{\Gamma})\in U_I} \min_{\upsilon\in\mathbb{R},\,\Lambda\in\mathbb{S}^n_+} \quad & tr(\Lambda\bar{\Gamma}) + \frac{\alpha}{1-\alpha}\upsilon - (r_f + (\bar{\mu} - r_f e)^\top x) \\ s.t. \quad & \\ & \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & \upsilon \end{bmatrix} \succeq 0. \end{aligned} \tag{37}$$

Convexity and compactness of feasible region and linearity of objective function with respect to $\bar{\mu}$ and $\bar{\Gamma}$ for fixed Λ and υ (and conversely) imply that we can exchange “min” and “max” to obtain

$$\begin{aligned} \min_{\upsilon\in\mathbb{R},\,\Lambda\in\mathbb{S}^n_+} \max_{(\bar{\mu},\bar{\Gamma})\in U_I} \quad & tr(\Lambda\bar{\Gamma}) + \frac{\alpha}{1-\alpha}\upsilon - (r_f + (\bar{\mu} - r_f e)^\top x) \\ s.t. \quad & \\ & \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & \upsilon \end{bmatrix} \succeq 0. \end{aligned} \tag{38}$$

Decompose now the inner maximizations into easier subproblems by applying the $U_{I_1} \times U_{I_2}$ representation of U_I to derive the following formulation of (38):

$$\begin{aligned} \min_{\upsilon\in\mathbb{R},\,\Lambda\in\mathbb{S}^n_+} \quad & -r_f + \frac{\alpha}{1-\alpha}\upsilon + \max_{\bar{\Gamma}\in U_{I_2}} tr(\Lambda\bar{\Gamma}) + \max_{\bar{\mu}\in U_{I_1}} -(\bar{\mu} - r_f e)^\top x \\ s.t. \quad & \\ & \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & \upsilon \end{bmatrix} \succeq 0. \end{aligned} \tag{39}$$

The dual formulations of the maximization problems in (39) are

$$\max_{\bar{\Gamma}\in U_{I_2}} tr(\Lambda\bar{\Gamma}) = \min_{\Lambda_+,\,\Lambda_-\succeq 0,\,\Lambda\preceq\Lambda_+-\Lambda_-} tr(\Lambda_+\bar{\Gamma}_+) - tr(\Lambda_-\bar{\Gamma}_-),$$

and

$$\max_{\bar{\mu}\in U_{I_1}} -(\bar{\mu} - r_f e)^\top x = \min_{x_+,\,x_-\geq 0,\,x=x_--x_+} (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_-.$$

Under suitable conditions —primal and dual strict feasibility— the duality gap in the first optimization problem above is zero, and we obtain the objective function of (35) as

$$\begin{aligned} \min_{\upsilon\in\mathbb{R},\,x_+,x_-\in\mathbb{R}^n,\Lambda,\,\Lambda_+,\,\Lambda_-\in\mathbb{S}^n} \quad & \frac{\alpha}{1-\alpha}\upsilon + tr(\Lambda_+\bar{\Gamma}_+) - tr(\Lambda_-\bar{\Gamma}_-) + \\ & (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_- \\ s.t. \quad & \\ & \begin{bmatrix} \Lambda & \frac{x_--x_+}{2} \\ \frac{(x_--x_+)}{2}^\top & \upsilon \end{bmatrix} \succeq 0, \\ & \Lambda \preceq \Lambda_+ - \Lambda_-, \\ & x_+, x_- \geq 0,\ \Lambda,\ \Lambda_+,\ \Lambda_- \succeq 0. \end{aligned} \tag{40}$$

To complete the robust counterpart (35), we need an explicit formulation of the robust counterpart of minimum return constraint:

$$\min_{(\bar{\mu},\bar{\Gamma})\in U_I} r_f + (\bar{\mu} - r_f e)^\top x \geq d.$$

To do this, we write the minimization problem as $-r_f + \max_{\bar{\mu}\in U_{I_1}} -(\bar{\mu} - r_f e)^\top x$ and use the result on dual form discussed above. Hence, the robust counterpart of minimum return constraint is:

$$r_f - (\bar{\mu}_+ - r_f e)^\top x_+ + (\bar{\mu}_- - r_f e)^\top x_- \geq d.$$

This completes the proof.

B Appendix: Data

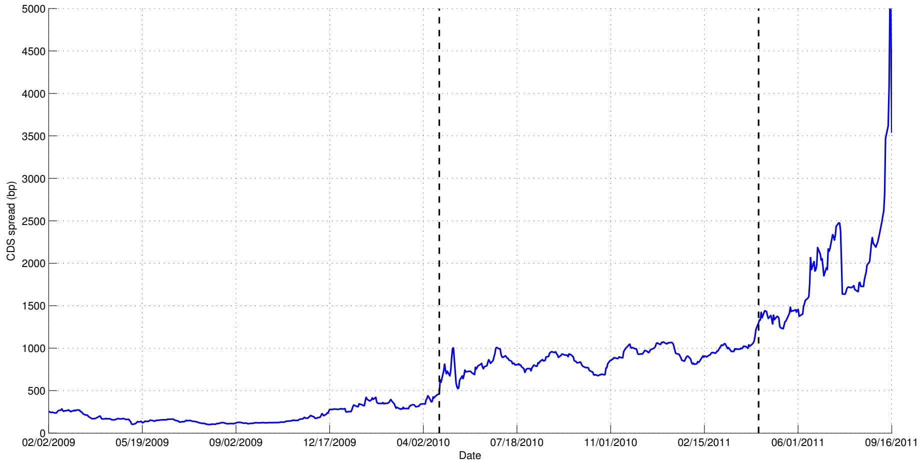


Fig. 13: CDS spreads of the Greek sovereign with identified regimes: Leftmost is the tranquil period, center is the turbulent, and the rightmost is the crisis.

	Stock.USA	Stock.UK	Stock.DE	Stock.JP	Bnd1.USA	Bnd2.USA
Mean	0.01296	0.01047	0.01057	-0.00189	0.00553	0.00702
SD	0.04101	0.04150	0.05796	0.06184	0.00467	0.01620
Skewness	-0.47903	-0.19051	-0.47281	0.04768	-0.18341	-0.07482
Kurtosis	3.76519	3.11399	4.11970	3.62119	2.77801	3.23974

Table 2: Moments of monthly differentials of the historical market data.

C Appendix: Hypothesis testing

We used the test of Wright et al. (2014) to test for the equality of Sharpe ratios of the returns created by the portfolios of pairs of models. The null hypothesis is as follows

$$H0 : \text{SH(Model1)} = \text{SH(Model2)}, \quad (41)$$

where SH(Model1) and SH(Model 2) are the Sharpe ratios of the sample of portfolio returns generated using Model1 and Model2, respectively. The null hypothesis is rejected if the test statistic $\hat{\chi}^2$ is greater than the critical value coming from the table of probabilities for the χ^2 distribution with one degree of freedom, at a pre-scpecified significance level. The p-value is the probability that, under the null hypothesis H0, the χ^2 value will be greater than the empirically estimated value $\hat{\chi}^2$, i.e. p-value = Prob $(\chi^2 > \hat{\chi}^2 \mid H0)$. The results of the test for different set of experiments are in Table 3 at the 0.01 significance level.

	Period	H0 hypothesis	$\hat{\chi}^2$	p-value
Active management eight-country sample	Turbulent-to-crisis	SH(VaR)=SH(RVaR)	0.0413	0.8389
		SH(CVaR)=SH(RCVaR)	0.0380	0.8454
	Crisis-to-post-crisis	SH(VaR)=SH(RVaR)	0.5874	0.4434
		SH(CVaR)=SH(RCVaR)	1.9031	0.1677
Active management eighteen-country sample	Turbulent-to-crisis	SH(CVaR)=SH(RCVaR)	0.7525	0.3857
	Crisis-to post-crisis	SH(CVaR)=SH(RCVaR)	0.0035	0.9529

Table 3: The hypothesis that the Sharpe ratios for different strategies are identical can not be rejected at the 0.01 significance level. The critical value for the χ^2 distribution with one degree of freedom is 6.6349.